

AFIT/DS/ENY/97-3

Control of Nonlinear Systems via State Feedback  
State-Dependent Riccati Equation Techniques

DISSERTATION

Kelly D. Hammett  
Captain, USAF

AFIT/DS/ENY/97-3

DTIC QUALITY INSPECTED 3

Approved for public release; distribution unlimited

19970708 118

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the Department of Defense or the U. S. Government.

DTIC QUALITY INSPECTED 8

AFIT/DS/ENY/97-3

Control of Nonlinear Systems via State Feedback  
State-Dependent Riccati Equation Techniques

DISSERTATION

Presented to the Faculty of the Graduate School of Engineering  
of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the  
Requirements for the Degree of  
Doctor of Philosophy

Kelly D. Hammett, B.S., M.S.

Captain, USAF

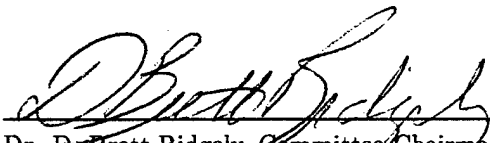
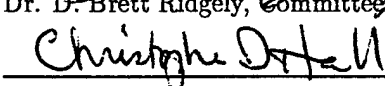
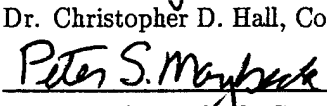
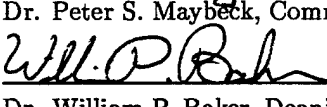
June, 1997

Approved for public release; distribution unlimited


Control of Nonlinear Systems via State Feedback  
State-Dependent Riccati Equation Techniques

Kelly D. Hammett, B.S., M.S.  
Captain, USAF

Approved:

 Dr. D. Brett Ridgely, Committee Chairman	<u>1 May 97</u>
 Dr. Christopher D. Hall, Committee Member	<u>1 May 97</u>
 Dr. Peter S. Maybeck, Committee Member	<u>1 May 97</u>
 Dr. William P. Baker, Dean's Representative	<u>1 May 97</u>

Accepted:

  
Dr. Robert A. Calico, Jr  
Dean, College of Engineering

## *Acknowledgements*

I would first of all like to thank God and my Lord Jesus Christ for the time, talents, and resources to carry out this adventure in higher learning. While it is true that insights often come as a result of hard work, inspiration, which I believe comes from the divine, plays a key role which cannot be underestimated. I also thank God for granting me a balanced perspective on life. I count each new day as a blessing not to be wasted, and thank God that He has allowed me to prioritize my life so that I may pursue stimulating work, while at the same time keeping my commitments to family and others foremost in my mind.

I would like to thank my wife, Kellie, for her continuous support of my academic studies, and, more importantly, for all the practical manifestations of that support, which gave me the freedom I needed to work when work was necessary. I also thank my children, Taylor and Chase, for being a source of inspiration and renewed joy each time I am around them.

I enjoyed a wonderful working relationship with my Research Advisor, Dr. D. Brett Ridgely, and with my other committee members, Dr. Chris Hall and Dr. Peter Maybeck. I thank you all for always being willing to spend time with me when I had a burning question or a hot new result to share. I thank you also for your excellence in your respective academic areas, which facilitated my learning and spurred me on to setting ever higher goals.

I thank Dr. Marc Jacobs of the Air Force Office of Scientific Research for sponsoring my work, and Dr. James Cloutier and Lt Col Curt Mracek of Wright Lab Armament Directorate for providing the initial motivation for my work, as well as continued feedback and discussion. I thank Mr. John Bowlus of the Flight Dynamics Directorate for the initial encouragement to apply for the AFIT PhD program, even though it meant I would be leaving his office prematurely.

I would like to thank my parents, Jerry and Martha Hammett, for always encouraging me and letting me try new things. You have always been supportive in a tangible way, and I appreciate all you have given me.

Finally, I would like to express gratitude to all my classmates and colleagues, for enlightening and often amusing conversations and technical exchanges. Having other people to share the journey with makes the road more pleasant. Thanks also to my PromiseKeepers small group, for continual encouragement and prayer support.

Kelly D. Hammett

# Table of Contents

	Page
Acknowledgements . . . . .	iii
List of Figures . . . . .	ix
Abstract . . . . .	x
 I. Introduction . . . . .	 1-1
1.1 Background . . . . .	1-1
1.2 Research Objectives and Contributions . . . . .	1-5
1.2.1 General Research Objectives . . . . .	1-6
1.2.2 SDRE Method-Specific Research Objectives . . . . .	1-6
1.3 ‘Linear’ Control of Nonlinear Systems - A Literature Review . . . . .	1-7
1.4 Overview . . . . .	1-9
 II. Background Theory . . . . .	 2-1
2.1 Stability Analysis of Nonlinear Dynamic Systems . . . . .	2-1
2.1.1 Lyapunov Based Concepts . . . . .	2-1
2.1.2 Stability in the First Approximation and Center Manifold Theory . . . . .	2-7
2.1.3 Useful Theorems and Lemmas . . . . .	2-12
2.2 Feedback Linearization . . . . .	2-14
2.2.1 Terms and Definitions from Differential Geometry . . . . .	2-15
2.2.2 Theory for Single-Input Single-Output Systems . . . . .	2-18
2.2.3 Extensions to Multi-Input Multi-Output Systems . . . . .	2-29
2.3 Recursive Backstepping . . . . .	2-35
2.3.1 Theory for Single-Input Systems . . . . .	2-35
2.3.2 Theory for Multi-Input Systems . . . . .	2-40
2.4 Nonlinear $H_\infty$ Control . . . . .	2-41
2.4.1 Problem Definition . . . . .	2-42
2.4.2 Induced $L_2$ Gain Analysis and Dissipativity for Nonlinear Systems . . . . .	2-43
2.4.3 State Feedback . . . . .	2-45
2.5 State-Dependent Riccati Equation Techniques . . . . .	2-49
2.5.1 The Nonlinear Regulator Problem . . . . .	2-49
2.5.2 Nonlinear $H_\infty$ Control Via the State Feedback SDRE Method . . . . .	2-55

III.	Motivational Example Problem . . . . .	3-1
	3.1 Problem Setup and Introductory Analysis . . . . .	3-1
	3.2 Feedback Linearization . . . . .	3-5
	3.3 Recursive Backstepping . . . . .	3-12
	3.4 Nonlinear $H_\infty$ Control . . . . .	3-17
	3.5 SDRE Techniques . . . . .	3-20
	3.5.1 Nonlinear Regulation Via the SDRE Method . . . . .	3-20
	3.5.2 Nonlinear $H_\infty$ Control Via the SDRE Method . . . . .	3-24
IV.	Insights into Design Issues . . . . .	4-1
	4.1 Partial Derivatives of Vector Matrix Products with $x$ Dependency . . . . .	4-1
	4.2 Convexity Concerns in SDRE Nonlinear Regulation . . . . .	4-4
	4.3 Symmetry of $P$ in Proposed HJI Solution $V = x^T P x$ . . . . .	4-6
	4.4 Simplification of SDRE Necessary Condition for Optimality . . . . .	4-8
	4.5 Relationships Between Solution Methods for SDRE Nonlinear $H_\infty$ Control . . . . .	4-11
	4.6 Optimal Control Numerical Solution Algorithm . . . . .	4-13
	4.7 Solvability of HJIs and HJB Equations . . . . .	4-19
V.	Solution Properties for Scalar Analytic Systems . . . . .	5-1
	5.1 Introduction . . . . .	5-1
	5.2 SDRE Solutions . . . . .	5-2
	5.3 Discussion . . . . .	5-6
	5.4 Examples . . . . .	5-7
	5.5 Conclusion . . . . .	5-9
VI.	Controllability Issues in SDRE Control . . . . .	6-1
	6.1 Introduction . . . . .	6-1
	6.2 Factored versus True Controllability of Nonlinear Systems . . . . .	6-2
	6.3 Theorems and Examples . . . . .	6-5
	6.4 Summary and Conclusions . . . . .	6-14
VII.	Stability of Systems with Full Rank Constant $B$ Matrices . . . . .	7-1
	7.1 Introduction . . . . .	7-1
	7.2 Control Algorithm . . . . .	7-1
	7.3 Global Asymptotic Stability via the Direct Method of Lyapunov . . . . .	7-2
	7.4 Discussion . . . . .	7-9
	7.5 Conclusion . . . . .	7-10

VIII.	Lyapunov Stability of Analytic Sampled Data Systems with Positive Definite $Q$ Matrices	8-1
8.1	Introduction	8-1
8.2	Sampled Data Nonlinear SDRE Regulation	8-2
8.3	Linear Transition Matrix Representation	8-3
8.4	Stability Via Lyapunov Theory	8-9
8.5	Discussion of Sampling Interval Size	8-14
8.6	Conclusion	8-16
IX.	Lyapunov Stability of Analytic Sampled Data Systems with Positive Semidefinite $Q$ Matrices	9-1
9.1	Introduction	9-1
9.2	Review of Stability Proof for $Q(x) > 0$	9-2
9.3	Stability for $Q(x) \geq 0$	9-3
9.3.1	LaSalle's Invariance Principle and Redefined Lyapunov Function	9-3
9.3.2	Characterization of $E$	9-5
9.3.3	Nonlinear Observability and Invariant Sets	9-6
9.3.4	Factored Observability and Control Action	9-10
9.3.5	Characterization of $M$	9-12
9.3.6	The Effect of Factored Controllability	9-14
9.3.7	Convergence of $P_k$ and $s_k \neq 0$	9-16
9.4	Examples	9-19
9.5	Summary and Conclusions	9-26
X.	Exponential Stability of SDRE Regulators	10-1
10.1	Introduction	10-1
10.2	Constant $F_k$ Matrices	10-1
10.3	Exponential Stability of SDRE	10-8
XI.	Extensions to SDRE Nonlinear $H_\infty$ Control	11-1
11.1	Standard Riccati Theory	11-1
11.2	Solution Properties of $H_\infty$ Scalar Analytic Systems	11-4
11.2.1	$H_\infty$ SDRE Solutions	11-5
11.2.2	Discussion	11-9
11.2.3	Examples	11-10
11.3	Lyapunov Stability for Systems with Full Rank, Constant $B$ Matrices	11-12
11.4	Asymptotic Stability of Sampled Data SDRE Nonlinear $H_\infty$ Controllers	11-15
11.5	Exponential Stability of Nonlinear SDRE $H_\infty$ Systems	11-17

XII.	Design Problem . . . . .	12-1
	12.1 Problem Description . . . . .	12-1
	12.2 Equations of Motion . . . . .	12-3
	12.3 Design Objectives . . . . .	12-5
	12.4 Open Loop System Analysis . . . . .	12-6
	12.5 Addition of Off-axis Rotor for Stabilizability/Detectability . . . . .	12-13
	12.6 SDRE Nonlinear Regulator Simulation Results . . . . .	12-21
	12.7 SDRE Nonlinear $H_\infty$ Simulation Results . . . . .	12-37
	12.8 Comparison with Other Methods . . . . .	12-40
	12.9 Conclusion . . . . .	12-44
XIII.	Conclusions and Recommendations for Further Research . . . . .	13-1
	13.1 Summary of Conducted Research . . . . .	13-1
	13.2 Conclusions . . . . .	13-4
	13.3 Recommendations for Further Research . . . . .	13-8
	Bibliography . . . . .	BIB-1
	Vita . . . . .	VITA-1

## *List of Figures*

Figure	Page
4.1 Optimal and Suboptimal State Histories . . . . .	4-17
4.2 Optimal and Suboptimal Control Histories . . . . .	4-17
4.3 Optimal $\alpha$ History . . . . .	4-18
4.4 Necessary Condition for Optimality History . . . . .	4-18
 9.1 State and Control Histories for $h = h_1$ . . . . .	 9-24
9.2 Histories of $V$ and $s$ for $h = h_1$ . . . . .	9-24
9.3 State and Control Histories for $h = h_2$ . . . . .	9-25
9.4 Histories of $V$ and $s$ for $h = h_2$ . . . . .	9-25
 12.1 Gyrostat Model of Axial Dual-Spin Spacecraft . . . . .	 12-4
12.2 SDRE State and Control Histories for 4-State Gyrostat ( $H = H_1$ ) . . . . .	12-23
12.3 SDRE State and Control Histories for 4-State Gyrostat ( $H = H_2$ ) . . . . .	12-23
12.4 State and Control Histories for 4-State Gyrostat ( $u = 0.02$ ) . . . . .	12-25
12.5 $x_1$ State Histories for 4-State Gyrostat . . . . .	12-25
12.6 Oblate SDRE State Histories for 5-State Gyrostat ( $H = H_1$ ) . . . . .	12-27
12.7 Oblate SDRE State Histories for 5-State Gyrostat ( $H = H_2$ ) . . . . .	12-27
12.8 Oblate State Histories for 5-State Gyrostat ( $u = 0.01$ ) . . . . .	12-28
12.9 Oblate Control Histories for 5-State Gyrostat . . . . .	12-29
12.10 Oblate $x_1$ Histories for 5-State Gyrostat . . . . .	12-29
12.11 Prolate State Histories for 5-State Gyrostat ( $u = 0.01$ ) . . . . .	12-30
12.12 Prolate SDRE State Histories for 5-State Gyrostat ( $H = H_1$ ) . . . . .	12-31
12.13 Prolate SDRE State Histories for 5-State Gyrostat ( $H = H_2$ ) . . . . .	12-31
12.14 Prolate Control Histories for 5-State Gyrostat . . . . .	12-32
12.15 Prolate $x_1$ Histories for 5-State Gyrostat . . . . .	12-32
12.16 Transverse SDRE State Histories for 5-State Gyrostat ( $H = H_1$ ) . . . . .	12-34
12.17 Transverse SDRE State Histories for 5-State Gyrostat ( $H = H_2$ ) . . . . .	12-34
12.18 Transverse SDRE State Histories for 5-State Gyrostat ( $H = H_2$ ) . . . . .	12-35
12.19 Transverse State Histories for 5-State Gyrostat ( $u = 0.01$ ) . . . . .	12-35
12.20 Transverse Control Histories for 5-State Gyrostat . . . . .	12-36
12.21 Transverse $x_1$ Histories for 5-State Gyrostat . . . . .	12-36
12.22 SDRE Nonlinear $H_\infty$ and Regulator Histories for Oblate 5-State Gyrostat ( $H = H_1$ )	12-39
12.23 SDRE Nonlinear $H_\infty$ and Regulator Controls for Oblate 5-State Gyrostat ( $H = H_2$ )	12-39

*Abstract*

Nonlinear regulation and nonlinear  $H_\infty$  control via state-dependent Riccati equation (SDRE) techniques are considered. Relationships between Hamilton-Jacobi/Bellman inequalities/equations and SDREs are examined, and a necessary condition for existence of solutions involving nonlinear stabilizability is derived. A single additional necessary criterion is given for the SDRE methods to yield the optimal control or guaranteed induced  $L_2$  gain properties. Pointwise stabilizability and detectability of factorizations prove necessary and sufficient, respectively, for well-posedness of standard numerical implementations of suboptimal SDRE regulators, but neither proves necessary if analytical solutions are allowed. For scalar analytic systems or those with full rank constant control input matrices, stabilizability and nonsingularity of the state weighting matrix function result in local and global asymptotic stability, respectively, due to equivalence between nonlinear and factored controllability in these cases. A proof of asymptotic stability for sampled data analytic SDRE controllers is also given, but restrictive assumptions make the main utility of these results guidance in choosing appropriate system factorizations. Conditions for exponential stability are also derived. All results are extendable to SDRE nonlinear  $H_\infty$  control with additional assumptions. The SDRE theory is illustrated by application to momentum control of a dual-spin satellite and comparison with other current methods.

# *I. Introduction*

## *1.1 Background*

Over the past several years, the relative maturity of linear controller design techniques, coupled with the inability of these techniques to handle strongly nonlinear dynamic systems satisfactorily, has led to a push in the development of nonlinear controller synthesis theory and techniques. As a result, theory has emerged for design according to a number of methods, including feedback linearization [33, 52, 62], variable structure control [21, 62], control Lyapunov functions [2, 64], recursive backstepping [40] and nonlinear  $H_\infty$  control [3, 34, 66]. Also, a state-dependent Riccati equation (SDRE) technique for nonlinear control has recently appeared in the literature [13, 14, 55]. Each method is currently at a different state of maturity, and each contains inherent limitations which restrict applicability to any given problem. A brief discussion of each method follows.

Feedback linearization is one of the more mature nonlinear controller synthesis techniques, having its basis in differential geometry, and receiving comprehensive treatment in several texts [33, 52, 62]. The basic idea is to find a coordinate transformation and static feedback which renders the closed loop system linear and stabilizable, and then to employ the wealth of existing linear controller synthesis techniques to the transformed system to achieve desired performance objectives. Theory exists primarily for state feedback but also for output feedback cases, and formulas for analytical solutions to stability, tracking, disturbance decoupling, and noninteracting control problems exist which may be solved, provided the system under consideration is feedback linearizable. The difficulties with this method stem from the fact that a given input-output system either possesses the desired feedback linearizability property or it does not. If the system is completely linearizable (all of the state equations may be made linear), then design may proceed. If the system is not completely linearizable and if the nonlinearized (unobservable) zero dynamics are asymptotically stable, controller design may proceed on the linearized portion of the state. However, if the zero dynamics are unstable, this method may not, in general, be used. Tests exist [33, 62] for checking the feedback linearizability

of a system and the stability of the resulting zero dynamics, and if a system is found not to pass the tests, this method is not applicable. An additional disadvantage of this method is that it completely removes all nonlinearities from a system, whether they are beneficial (e.g., with regard to stability) or not. Thus, the method is nonoptimal. Finally, technically only a number of outputs equal to the number of inputs can be controlled, posing an additional design constraint. For cases where there are more inputs than outputs, pseudo-controls can be formed. Several examples of application of this method to suitable systems can be found in the literature [43, 61, 63], and current research in this area focuses on extending the applicability of the method to systems which currently may not be addressed (systems not completely linearizable or with unstable zero dynamics). Also, extending other differential geometry concepts to achieve more advanced performance objectives is being pursued.

Variable structure control (VSC), also known as sliding mode control [62], is another relatively mature nonlinear controller synthesis method. The main objective is to define a manifold in which it is desired to keep the system state, and then to drive the system to this manifold (i.e., make the manifold attractive in the closed loop). Variable structure control results in discontinuous, high-gain controllers, with ensuing chattering of the control. The main advantage of VSC is that its high-gain nature yields desirable robustness properties under certain conditions, making VSC a good candidate for control of systems with appropriately modeled uncertainty [62]. Current research is attempting to eliminate the chattering of the controls, and seeks optimal definitions for the desired state space manifold.

The idea behind control Lyapunov functions is to try to find a Lyapunov function [68] for the closed loop system and to derive a controller which guarantees stability, based on application of the Direct Lyapunov Method [36] to the closed loop system equations. Necessary and sufficient conditions have been derived for existence of control Lyapunov functions for certain classes of problems [2], but as yet no systematic method for finding them has been proposed. Thus, this method is immature in terms of synthesis capability.

Recursive backstepping is another relatively new technique being advocated as an extremely flexible yet systematic methodology for direct nonlinear control design. A recent text gives a comprehensive treatment of the method [40], claiming that its primary advantage lies in its ability to address adaptive control of nonlinear systems, thus going one step beyond simple nominal control design. The basic idea behind this method is sequential scalar control design for each nonlinear state differential equation, incorporating the use of subsequent states or combinations of states as pseudocontrols in the previous designs. The actual physical control is made to appear in the final state equation, and thus is the last quantity chosen to attempt to meet the overall design objectives. The method is Lyapunov function based, and thus aims to guarantee closed loop stability. The designer attempts to obtain other performance objectives by “optimizing” coefficients in the Lyapunov function, and by clever choice of pseudocontrols. From this discussion, it can be seen that recursive backstepping is really an application and extension of control Lyapunov function theory [40]. Recursive backstepping, as with other methods, is limited in the types of systems to which it can be applied, and although the general theory is systematic, the actual synthesis details are problem- and designer-specific. Thus, the method can yield several different designs depending on the designer’s insight and the particular choices made.

Nonlinear  $H_\infty$  control theory is currently in the early stages of development [3, 4, 31, 34, 35, 66, 67, 75]. The current state-of-the-art consists of sufficient conditions for obtaining nonlinear controllers which guarantee induced  $L_2$  performance and local closed loop stability for both state and output feedback cases. In the state feedback case, the sufficient condition corresponds to finding a locally positive definite solution to a certain Hamilton-Jacobi inequality (HJI) or Hamilton-Jacobi-Isaacs equation (HJIE) (or finding a positive semidefinite solution with additional assumptions required), which depends parametrically on the induced  $L_2$  gain. Solving the output feedback problem requires solution of another HJI and ensuring satisfaction of a third condition relating the solutions of the two HJIs. Nontrivial solutions have been found for limited special cases, mainly involving lossless/dissipative systems and state feedback [15, 16, 37, 48, 71]. For the general case, however,

although initial (numerical) solution attempts have been made [29], little information exists in the literature describing how solutions to HJIs can be found, let alone guaranteeing satisfaction of the HJI solution coupling condition of the output feedback problem. Also, it remains to discover a way to compute the optimal induced  $L_2$  gain, although recently methods for its approximation have appeared in the literature [35, 75]. Thus, much work remains to be done in characterizing and actually obtaining solutions to the nonlinear  $H_\infty$  synthesis problem.

Finally, a state-dependent Riccati equation (SDRE) based technique for nonlinear regulation, originally proposed as far back as 1962 [54], has recently seen renewed interest [13, 14, 55]. This method has been described as a nonlinear extension of the well known two-Riccati-equation solution technique to linear quadratic Gaussian (LQG) type synthesis problems. SDRE solution techniques for both state and output feedback nonlinear  $H_\infty$  control have also been proposed, but not developed. Even though variants of this method have been proposed for a number of years [9, 19, 54, 70], the underlying theory is scant, despite the fact that results from its application have been impressive [14, 54]. Several academic problems have been solved analytically, and a numerically solved output feedback SDRE regulator solution has been obtained for a nonlinear missile autopilot [12], indicating the feasibility of applying this method to realistic problems. Also, a thesis at AFIT [53] was recently completed applying this method to satellite control and control of an artificial human pancreas. A tremendous need exists, however, to fill in a large number of theoretical 'holes' for this method. Stability, optimality, factorization/parametrization techniques, solution methods and a number of other issues still need to be analyzed in detail.

We close this section with the observation that the above discussion attempts to focus on nonlinear control techniques which have emerged relatively recently, and thus, the long-established optimal control techniques based on variational calculus and/or dynamic programming have been intentionally omitted. These are certainly valid techniques for control design for nonlinear systems, but the theoretical basis for them is quite mature and well documented. See, for example [8]. Application of the theory generally results in nonlinear, constrained, two-point boundary value problems

(variational calculus) or partial differential (Bellman) equations (dynamic programming), analytical solutions to either of which are extremely difficult to obtain for most nontrivial problems of interest. Although some research continues in the area of control design by these methods, it mainly focuses on application of the relevant theory to limited special cases to obtain the form of the optimality equations, or on numerical attempts to solve them, and not on extending the theory itself. We mention them in passing here because some techniques examined in this document, particularly the SDRE nonlinear regulation method, may be viewed as new ways of solving special cases of such optimal control problems, and thus must yield solutions equivalent to those obtained from using the pertinent theory.

## *1.2 Research Objectives and Contributions*

As can be seen from the preceding discussion, a number of potential design techniques now exists for direct nonlinear control design. Although several methods are fairly well established theoretically, many of them have severe limitations which restrict their applicability or usefulness to realistic control design problems. This research focuses on providing a sound theoretical basis for control of nonlinear systems via the state feedback SDRE techniques, which, as discussed above, have proven quite successful in a number of simulated applications, but lack supporting theory in a number of areas. Many of these issues are attacked herein, with significant headway being made in theoretical justification for design choices and properties. During the course of our presentation, wherever possible we seek to draw comparisons between the SDRE techniques developed herein and other established nonlinear control techniques, particularly the methods of feedback linearization, recursive backstepping, Hamilton-Jacobi-Bellman optimal control, and nonlinear  $H_\infty$  control. To this end a summary of the theory relevant to these methods is included herein, and a motivational design example applying all these techniques to the same problem is pursued in detail. In addition to theoretical development, we also make a nontrivial application of the SDRE control techniques to a problem of interest to the U.S. Air Force. The particular control problem addressed is gyrostat

satellite angular momentum control [22, 23, 25]. This problem has characteristics which invite the use of direct nonlinear control techniques. The gyrostat problem exhibits strongly nonlinear dynamics: so much so that, in fact, the linearized dynamics for the baseline problem are zero. The lack of a stabilizable linearization for this problem poses a severe challenge to the control design techniques, requiring a configuration modification to obtain acceptable solutions. These types of issues are often ignored in purely academic studies of nonlinear control design, but are included herein in our examination of applicability of design methods. A summary of general and specific research objectives considered is given below. The remainder of this dissertation serves to document their satisfaction and illustrate the contributions of this research.

### *1.2.1 General Research Objectives*

- To investigate the solvability of state-of-the-art control synthesis algorithms for nonlinear dynamic systems of practical interest to the U.S. Air Force
- To develop methods for solving the above synthesis problems more easily
- To extend the applicability of current methods where possible
- To make an objective determination of the relative merits and disadvantages of the various techniques considered

### *1.2.2 SDRE Method-Specific Research Objectives*

- To investigate the relative tradeoffs between analytical/numerical solution of SDREs
- To analyze the effects of various state-dependent parametrizations
- To examine correlations with existing proposed nonlinear  $H_\infty$  solution techniques
- To identify and prove conditions leading to closed loop stability
- To investigate sampled data implementation issues

- To discover relationships between SDRE and optimal nonlinear regulation
- To seek correlations with the known linear Riccati,  $H_2$ , and  $H_\infty$  theory
- To apply the method to control design of a nontrivial nonlinear dynamic system
- To make an objective determination of the relative merits and disadvantages of the design technique

### 1.3 'Linear' Control of Nonlinear Systems - A Literature Review

When speaking strictly of SDRE techniques, the literature citations are few. However, over the past several years, a number of researchers have proposed nonlinear control algorithms which involve application of linear design methods to linear-like 'factored' representations of a nonlinear system [9, 13, 19, 21, 54, 55, 70]. For continuous time, state feedback, input-affine, autonomous nonlinear dynamic systems of the form

$$\begin{aligned} \dot{x} &= a(x) + b(x)u, \quad a(0) = 0 \\ z &= \begin{bmatrix} h(x) \\ \bar{R}(x)u \end{bmatrix}, \quad h(0) = 0 \end{aligned} \tag{1.1}$$

with state vector  $x \in \mathcal{R}^n$ , control vector  $u \in \mathcal{R}^m$ , penalized variable  $z \in \mathcal{R}^s$ , and nonsingular (for all  $x$ ) control penalty matrix function  $\bar{R}(x)$ , it is assumed that one can obtain a factored representation of the form

$$\begin{aligned} \dot{x} &= A(x)x + B(x)u \\ z &= \begin{bmatrix} H(x)x \\ \bar{R}(x)u \end{bmatrix} \end{aligned} \tag{1.2}$$

so that at each point in the state space the nonlinear vector fields  $a(x)$  and  $h(x)$  have a linear appearance given by  $A(x)$  and  $H(x)$ . This concept has alternatively been called 'equivalent linearization' [9, 54], 'apparent linearization' [70], 'extended linearization' [21], or most recently obtaining a 'state-dependent coefficient factorization' [13] for (1.1). In [54], a suboptimal approach to the finite time

optimal regulator problem for input-affine nonlinear systems is proposed, which requires solution of a state-dependent Riccati differential equation (RDE) involving the factorization (1.2). Several low-order example problems are solved analytically, suboptimal versus optimal performance comparisons are made, and global asymptotic stability is proven for the suboptimally-controlled systems. In [9], a variation of the above approach is proposed, in which the RDE is solved by power series expansion of the RDE solution with respect to the states about a user-selected fixed point, solving a nonstate-dependent RDE to obtain the zeroth-order solution term, and solving several linear matrix differential equations to obtain the higher-order coefficients in the power series expansion for the state-dependent RDE solution. In [70], a suboptimal solution for the nonlinear state feedback infinite time horizon quadratic regulation problem is proposed, which involves solving an algebraic state-dependent Riccati equation (SDRE) for any location traversed in the state space. In [19], the same basic idea is revisited, and conditions relating the suboptimal solution to the optimal solution are derived. In both these papers, the state and control weights  $\bar{R}$  and  $H$  are assumed to be constant matrices, so that the regulation problem is indeed quadratic. More recently, in [21] application of any linear control algorithm to (1.2) is suggested, but no theoretical justification for such an approach is given. In [13] both state and output feedback SDRE approaches to regulation and nonlinear  $H_\infty$  control problems are proposed, where weighting matrices are not restricted to be constants, but may instead be functions of  $x$ . Local stability is proven for suboptimal state feedback versions of the above, and an additional necessary condition which must be satisfied for optimality of the state feedback regulator is given. In [55] a Lyapunov function is proposed for establishing global stability of the suboptimal state feedback regulator, based on a restricted class of weighting matrix functions. Finally, in [30], the SDRE nonlinear regulator is examined as an alternative to solving Hamilton-Jacobi-Isaacs equations in the optimal regulation problem, and sufficient conditions for existence of an 'optimal' state-dependent factorization are given. While some progress has been made in theoretically justifying such 'linearized' methods, much remains to be done. Important areas such as analytical versus numerical solution techniques, guaranteed stability beyond the domain of

attraction of the system linearization, controllability issues, choices of factorizations, sampled data applications, and optimality remain to be fully investigated. It is precisely these issues which are addressed in this research.

## *1.4 Overview*

The remainder of this document is organized as follows. Chapter 2 summarizes the existing theory relevant to nonlinear stability analysis and control design via the four methods discussed above. Chapter 3 extends the presentation of material in Chapter 2 by showing application of each method to an academic second-order design example. This initial attempt at design gives an early indication of each method's utility, strengths, and weaknesses, and thereby points to promising specific areas of research. The next several chapters summarize the theoretical advances achieved by this research. Chapter 4 contains a number of insights into various design issues, in particular addressing convexity, optimality, and analytical versus numerical solution approach tradeoffs. In Chapter 5, we explore the properties of scalar analytic systems regulated by SDRE control, developing necessary and sufficient conditions for existence of locally stabilizing analytic solutions. In Chapter 6, we explore the correlations between true nonlinear system controllability and controllability of state-dependent factorizations, and explain the impact of each on system stability. We build on a special case of these results in Chapter 7, in which we prove global asymptotic stability of continuous time systems with full rank, constant  $B$  matrices and globally positive definite state weighting matrix functions. In the next two chapters we analyze sampled data implementations of SDRE nonlinear regulators, deriving sufficient conditions for semiglobal asymptotic stability of the closed loop system. Then, in Chapter 10, we explore some issues relating to exponential stability of SDRE regulators. In Chapter 11, we revisit the theory of Chapters 5-10, showing that it may be extended to the setting of nonlinear  $H_\infty$  control via the SDRE method, provided some suitable additional assumptions are made. In Chapter 12, the proposed satellite control design problems and dynamic models are presented, and results of design iterations and simulations are shown, illustrating the theory laid out in the previous

chapters. Finally, Chapter 13 concludes this dissertation by summarizing the conducted research and highlighting specific areas of contribution.

## II. Background Theory

This chapter presents background theory necessary for stability analysis and synthesis of nonlinear dynamic control systems. In Section 2.1, relevant analysis tools are presented, consisting primarily of fundamental stability concepts and theory. The remaining sections of Chapter 2 cover the basic requisite theory for direct synthesis of nonlinear control systems for the methods mentioned in Section 1.2, namely feedback linearization, recursive backstepping, nonlinear  $H_\infty$  control, and the SDRE methods, respectively.

### 2.1 Stability Analysis of Nonlinear Dynamic Systems

#### 2.1.1 Lyapunov Based Concepts

We now address fundamental terminology and theory needed for discussing stability of nonlinear dynamic systems. More precisely, the following discussion refers to solutions and equilibria of dynamic systems and not to the systems themselves. The following definitions are taken primarily from [40]. Let  $\mathcal{R}_+$  be the set of nonnegative real numbers, and consider a nonlinear, nonautonomous dynamic system with equations of motion of the form

$$\dot{x} = a(x, t) \tag{2.1}$$

with state  $x \in \mathcal{R}^n$ , and  $a : \mathcal{R}^n \times \mathcal{R}_+ \rightarrow \mathcal{R}^n$  piecewise continuous in  $t$  and locally Lipschitz in  $x$ . The solution of (2.1) which starts from the point  $x_0$  at time  $t_0 \geq 0$  is denoted as  $x(t; x_0, t_0)$  so that  $x(t_0; x_0, t_0) = x_0$ . Let  $\tilde{x}_0 \in \mathcal{R}^n$  be an equilibrium of (2.1), i.e.,  $a(\tilde{x}_0, t) = 0 \forall t \in \mathcal{R}_+$ , and let  $\|x\|$  represent the Euclidean norm of  $x \in \mathcal{R}^n$ . The following definitions will prove useful. We shall say that the (unique) solution  $x(t; x_0, t_0)$  of (2.1) is

- **bounded**, if  $\exists$  a constant  $B(x_0, t_0) > 0$  such that

$$\|x(t; x_0, t_0)\| < B(x_0, t_0) \forall t \geq t_0 \tag{2.2}$$

- **(Lyapunov) stable**, if for each  $\varepsilon > 0 \exists$  a  $\delta(\varepsilon, t_0) > 0$  such that

$$\|\tilde{x}_0 - x_0\| < \delta \Rightarrow \|x(t; \tilde{x}_0, t_0) - x(t; x_0, t_0)\| < \varepsilon \quad \forall t \geq t_0 \quad (2.3)$$

- **attractive**, if  $\exists$  an  $r(t_0) > 0$  and for each  $\varepsilon > 0 \exists$  a  $T(\varepsilon, t_0) > 0$  such that

$$\|\tilde{x}_0 - x_0\| < r \Rightarrow \|x(t; \tilde{x}_0, t_0) - x(t; x_0, t_0)\| < \varepsilon \quad \forall t \geq t_0 + T \quad (2.4)$$

- **asymptotically stable**, if it is both stable and attractive, so that if we define  $e_x(t) = \|x(t; \tilde{x}_0, t_0) - x(t; x_0, t_0)\|$ , the solution  $x(t; x_0, t_0)$  is said to be asymptotically stable if

$$\lim_{t \rightarrow \infty} e_x(t) = 0 \quad (2.5)$$

- **exponentially stable**, if asymptotic stability holds with an exponential upper bound on the rate of convergence, so that

$$e_x(t) \leq k e^{-\alpha t} e_x(0) \quad (2.6)$$

for some  $k > 0, \alpha > 0$ .

- **unstable**, if it is not stable.

If the stability properties of a solution are independent of  $t_0$ , they are said to be **uniform**. Thus, all time-invariant systems have uniform stability properties. All of the above terms except bounded can also be used to describe equilibria of (2.1). A stable or asymptotically stable equilibrium,  $\tilde{x}_0$ , has a **region of attraction** - a set  $\Omega$  of initial states  $x_0$  which either remain close to  $\tilde{x}_0$  in the sense of (2.3) (for  $\tilde{x}_0$  stable), or converge to  $\tilde{x}_0$  as in (2.5) (for  $\tilde{x}_0$  asymptotically stable). If  $\Omega = \mathcal{R}^n$ , the stability properties of  $\tilde{x}_0$  are said to be **global**. Note that the definitions of (2.3) and (2.4) require the existence of an open ball in  $\mathcal{R}^n$ ,  $B_n$ , centered at  $\tilde{x}_0$  and with radius  $\delta$  or  $r$ , respectively ( $B_n(\tilde{x}_0, \xi) \equiv \{x \in \mathcal{R}^n \mid \|x - \tilde{x}_0\| < \xi\}$ ), such that  $B_n \subseteq \Omega$ . Thus, a region of attraction extends at least some finite distance in *all* directions from a stable equilibrium.

As a matter of convenience, if we are interested in examining the stability of an equilibrium point, we will often translate coordinates by the relation  $z = x - \tilde{x}_0$ , so that in the new coordinate system we have

$$\dot{z} = \dot{x} = a(z + \hat{x}_0, t) \equiv \bar{a}(z, t) \quad (2.7)$$

implying

$$\bar{a}(0, t) = a(\hat{x}_0, t) = 0 \quad (2.8)$$

Thus, in the new coordinate system,  $z = 0$  is an equilibrium of (2.7). Without loss of generality we will therefore take the equilibrium in (2.1) to be the origin in what follows.

We are interested in ways of determining the stability of an equilibrium, without solving (2.1) directly. As a preliminary, we need a few more definitions. A function  $V : U \rightarrow \mathcal{R}_+$ , where  $U$  is some neighborhood of the origin (an open set containing  $x = 0$ ), is said to be

- **positive definite** ( $> 0$ ) on  $U$  iff  $V(0) = 0$  and  $V(x) > 0 \forall x \in U, x \neq 0$
- **positive semidefinite** ( $\geq 0$ ) on  $U$  iff  $V(0) = 0$  and  $V(x) \geq 0 \forall x \in U$
- **negative definite** ( $< 0$ ) on  $U$  iff  $V(0) = 0$  and  $V(x) < 0 \forall x \in U, x \neq 0$
- **negative semidefinite** ( $\leq 0$ ) on  $U$  iff  $V(0) = 0$  and  $V(x) \leq 0 \forall x \in U$
- **radially unbounded** if  $U = \mathcal{R}^n$  and  $V(x) \rightarrow \infty$  as  $\|x\| \rightarrow \infty$

Definiteness of a time-varying function,  $V(x, t)$ , is defined by requiring the existence of a continuous, nondecreasing, and definite function,  $W(x)$ , as defined above which bounds  $V(x, t)$  away from zero in the appropriate way. For example, for  $V(x, t)$  to be positive definite requires

$$W(0) = 0, V(x, t) \geq W(x) > 0 \forall x \in U, x \neq 0, \forall t \geq 0$$

Another useful concept in Lyapunov stability theory for time-varying functions is that of a function  $V(x, t)$  being **decreascent**, which implies the existence of a continuous, nondecreasing scalar function,  $Z$ , such that  $Z(0) = 0$  and, for all  $t \geq 0$ ,

$$V(x, t) \leq Z(\|x\|)$$

We now present theorems which accomplish the above-stated objective.

**Theorem 2.1.1 (Lyapunov)** *Let  $x = 0$  be an equilibrium of (2.1) and let  $U(0, r) \in \mathcal{R}^n$  be an open ball of radius  $r \in \mathcal{R}_+$  centered at the origin as defined above. Also let  $V : U \rightarrow \mathcal{R}_+$  be a continuously differentiable positive definite function  $V(x)$  on  $U$  such that*

$$\dot{V} = \frac{\partial V}{\partial x} a(x, t) \leq 0 \quad \forall t \geq 0, \forall x \in U \quad (2.9)$$

*Then the origin is a locally stable equilibrium of (2.1). If  $\dot{V}$  in (2.9) is strictly less than zero, then the origin is a locally asymptotically stable (LAS) equilibrium of (2.1).*

**Proof:** See [68], Theorems 8.1, 8.2. ■

A function  $V$  satisfying all the conditions given in Theorem 2.1.1 is often called a **Lyapunov function** for (2.1). The above basic stability theorem only provides local results. The following theorem allows us to obtain global stability results.

**Theorem 2.1.2 (Lasalle-Yoshizawa)** *Let  $x = 0$  be an equilibrium of (2.1) and let  $V : \mathcal{R}^n \rightarrow \mathcal{R}_+$  be a continuously differentiable, positive definite, and radially unbounded function  $V(x)$  such that*

$$\dot{V} = \frac{\partial V}{\partial x} a(x, t) \leq -W(x) \leq 0 \quad \forall t \geq 0, \forall x \in \mathcal{R}^n \quad (2.10)$$

*where  $W$  is a continuous function. Then, all solutions of (2.1) are globally uniformly bounded and satisfy*

$$\lim_{t \rightarrow \infty} W(x(t)) = 0 \quad (2.11)$$

*In addition, if  $W(x)$  is positive definite, then the origin is globally, uniformly asymptotically stable.*

**Proof:** See [40], Theorem A.8. ■

A converse form of Theorem 2.1.1 exists which allows us to prove instability of an equilibrium.

**Theorem 2.1.3 (Converse Lyapunov)** *Let  $x = 0$  be an equilibrium of (2.1) and let  $U(0, r) \in \mathcal{R}^n$  be an open ball of radius  $r \in \mathcal{R}_+$  centered at the origin as defined above. Also let  $V : U \rightarrow \mathcal{R}_+$  be a continuously differentiable positive definite function  $V(x)$  on  $U$ , and  $W(x)$  continuous and positive definite on  $U$  such that*

$$\dot{V} = \frac{\partial V}{\partial x} a(x, t) \geq W(x) > 0 \quad \forall t \geq 0, \forall x \text{ in some neighborhood } N \text{ of the origin} \quad (2.12)$$

*Then the origin is an unstable equilibrium of (2.1).*

**Proof:** See [68], Theorem 8.3. ■

If we want to consider only time-invariant nonlinear systems

$$\dot{x} = a(x) \tag{2.13}$$

then LaSalle's Invariance Theorem and its associated asymptotic stability theorem will also prove useful. A prerequisite is the definition of an **invariant set**. A set  $M \subseteq \mathcal{R}^n$  is said to be invariant with respect to (2.13) if, for some  $t_0 \geq 0$ ,

$$x(t_0) = x_0 \in M \Rightarrow x(t; x_0, t_0) \in M \quad \forall t \in \mathcal{R}_+ \tag{2.14}$$

while  $M$  is said to be **positively invariant** if

$$x(t_0) = x_0 \in M \Rightarrow x(t; x_0, t_0) \in M \quad \forall t \geq t_0 \tag{2.15}$$

We now present the theorems.

**Theorem 2.1.4 (Lasalle Invariance)** *Let  $\Omega$  be a positively invariant set of (2.13) and let  $V : \Omega \rightarrow \mathcal{R}_+$  be a continuously differentiable function  $V(x)$  such that  $\dot{V}(x) \leq 0 \quad \forall x \in \Omega$ . Also, let  $E = \{x \in \Omega \mid \dot{V}(x) = 0\}$ , and let  $M$  be the largest invariant set contained in  $E$ . Then every bounded solution  $x(t)$  starting in  $\Omega$  converges to  $M$  as  $t \rightarrow \infty$ .*

**Proof:** See [44]. ■

**Theorem 2.1.5 (Time-invariant GAS)** *Let  $x = 0$  be the only equilibrium of (2.13) and let  $V : \mathcal{R}^n \rightarrow \mathcal{R}_+$  be a continuously differentiable, positive definite, radially unbounded function  $V(x)$  such that  $\dot{V}(x) \leq 0 \quad \forall x \in \mathcal{R}^n$ . Also, let  $E = \{x \in \mathcal{R}^n \mid \dot{V}(x) = 0\}$ , and suppose no solution other than  $x(t) = 0$  can stay forever in  $E$ . Then the origin is globally asymptotically stable (GAS).*

**Proof:** See [44]. ■

In later sections of this dissertation, we will want to consider stability of systems with inputs. Thus, we need a final stability concept, introduced by Sontag [65]. A system

$$\dot{x} = f(x, u) \quad (2.16)$$

is said to be **input-to-state stable** (ISS) if for any  $x(0)$  and for any input  $u$  continuous and bounded on  $[0, \infty)$  the solution to (2.16) exists  $\forall t \geq 0$  and satisfies

$$\|x(t)\| \leq \beta(\|x(0)\|, t) + \gamma \left( \sup_{0 \leq \tau \leq t} \|u(\tau)\| \right), \quad \forall t \geq 0 \quad (2.17)$$

where  $\beta(s, t)$  and  $\gamma(s)$  are strictly increasing functions of  $s \in \mathcal{R}_+$  with  $\beta(0, t) = 0, \gamma(0) = 0$ , while  $\beta$  is a decreasing function of  $t$  with  $\beta \rightarrow 0$  as  $t \rightarrow \infty, \forall s \in \mathcal{R}_+$ . This definition is appropriate for nonlinear systems in that it includes effects of the control and of the initial condition. The ISS nature of a system (2.16) can be established by at least two methods other than simply using the definition: via a closed loop Lyapunov type argument and via a dissipation inequality argument. These are presented in the following theorem, but first we present some useful definitions.

A continuous function  $\rho : [0, a) \rightarrow \mathcal{R}_+$  is said to be of class  $\mathcal{K}$  if it is positive definite and strictly increasing. It is said to be of class  $\mathcal{K}_\infty$  if, in addition,  $a = \infty$  and it is radially unbounded.

**Theorem 2.1.6** *Suppose that for (2.16)  $\exists$  a continuously differentiable class  $\mathcal{K}_\infty$  function  $V : \mathcal{R}^n \times \mathcal{R}_+ \rightarrow \mathcal{R}_+$ , class  $\mathcal{K}_\infty$  functions  $\kappa, \rho_1$  and  $\rho_2$ , and a class  $\mathcal{K}$  function  $\psi$  such that*

*i.  $\forall x \in \mathcal{R}^n$  and  $u \in \mathcal{R}^m$ ,*

$$\|x\| \geq \kappa(\|u\|) \Rightarrow \frac{\partial V}{\partial x} f(x, u) \leq -\psi(\|x\|) \quad (2.18)$$

*or,*

*ii.*

$$\frac{\partial V}{\partial x} f(x, u) \leq -\rho_1(\|x\|) + \rho_2(\|u\|) \quad (2.19)$$

*Then (2.16) is ISS, and if item i above holds then  $\gamma$  in (2.17) is equal to  $\kappa$ , and we call  $V$  in (2.18) an ISS-Lyapunov function.*

### 2.1.2 Stability in the First Approximation and Center Manifold Theory

In the latter part of Section 2.1.1, we presented two theorems based on Lyapunov arguments that are useful for determining the stability of equilibria of nonlinear time invariant systems described by (2.13). Another well known test for determining local stability of equilibria of (2.13) is the so-called *Principle of Stability in the First Approximation*. If we assume that  $a$  in (2.13) is at least twice continuously differentiable and that the equilibrium of interest is again the origin, then the following statements can be made about the local asymptotic stability of (2.13) based on the linear approximation of  $a$  at  $x = 0$  [68]. Let

$$J_0 = \left[ \frac{\partial a}{\partial x} \right]_{x=0} \quad (2.20)$$

denote the Jacobian matrix of  $a$  at  $x = 0$ . Then if

- all the eigenvalues of  $J_0$  have negative real parts, the origin is a locally asymptotically stable equilibrium of (2.13),
- at least one eigenvalue of  $J_0$  has a positive real part, the origin is an unstable equilibrium of (2.13).

The *Principle of Stability in the First Approximation* obviously does not cover all cases of interest. In particular, it provides no information when all of the real parts of the eigenvalues of  $J_0$  are nonpositive, and at least one eigenvalue has a zero real part. When this is the case, Center Manifold Theory may often be used to draw conclusions regarding the local stability properties of an equilibrium point for a time-invariant system. We point out that entire texts have been written on this topic, and in order to keep the level of detail manageable, we will present just enough of the theory to meet our analysis purposes. This material is extracted from [33], Appendices A and B, to which the reader is referred for a more comprehensive treatment of the subject. We will again need a few definitions before proceeding.

Recall that an **open set**  $O$  in  $\mathcal{R}^n$  is defined as a collection of points  $x \in \mathcal{R}^n$  such that for every  $x \in O$  an open ball,  $B(x, r)$ , exists which is wholly contained in  $O$  [51]. Unless stated otherwise, when we use the term **smooth manifold**, we are simply referring to some special open subset of  $\mathcal{R}^n$ . In fact, any open subset of  $\mathcal{R}^n$  is a smooth manifold of dimension  $n$  [33]. With this in mind, our earlier definition of an invariant set readily extends in the expected way to include the term invariant manifold. However, to define a center manifold, we need the concept of a locally invariant manifold. A manifold  $N$  is said to be **locally invariant** with respect to (2.13) if for each  $x_0 \in N$ ,  $\exists t_1 > 0$  such that  $x(t; x_0, 0) \in N \forall t \in (0, t_1)$  (note that we have arbitrarily set  $t_0 = 0$ , which we can do without loss of generality since (2.13) is time-invariant). Another required definition is that of the tangent space of a smooth manifold at a point. This concept requires some development, presented below.

Let  $N$  be a smooth manifold of dimension  $n$ . The vector space of all functions defined on  $\mathcal{R}^n$  that are  $r$  times continuously differentiable we shall denote by  $C^r$ . We shall say a function is **smooth** if it is infinitely times continuously differentiable. Thus, the vector space of all functions defined on  $\mathcal{R}^n$  that are infinitely times continuously differentiable we shall denote by  $C^\infty$ . Now, let  $p$  be any point in  $N$ . A real-valued function  $\lambda$  is said to be **smooth in a neighborhood of  $p$**  if the domain of  $\lambda$  includes an open subset  $U$  of  $N$  containing  $p$  and the restriction of  $\lambda$  to  $U$  is a smooth function as defined above. We denote the set of all functions smooth in a neighborhood of  $p$  by  $C^\infty(p)$ , noting that  $C^\infty(p)$  is a vector (linear) space over the field  $\mathcal{R}$ , and in fact  $C^\infty(p)$  is a commutative linear algebra since multiplication of vectors is well-defined and order-independent. We now give the definition of a tangent vector at a point in a smooth manifold.

A **tangent vector**  $v$  at  $p$  is a map  $v : C^\infty(p) \rightarrow \mathcal{R}$  satisfying the following two properties:

- i. (linearity):  $v(a\lambda + b\gamma) = av(\lambda) + bv(\gamma) \forall \lambda, \gamma \in C^\infty(p) \text{ and } a, b \in \mathcal{R}$
- ii. (Leibniz Rule):  $v(\lambda\gamma) = \gamma(p)v(\lambda) + \lambda(p)v(\gamma) \forall \lambda, \gamma \in C^\infty(p)$

We define the **tangent space** to  $N$  at  $p$ , written  $T_p N$ , as the set of all tangent vectors at  $p$ . It is simple to show that the set  $T_p N$  forms a vector space over  $\mathcal{R}$ , with the standard rules of vector

addition and scalar multiplication. The tangent space to a smooth manifold  $N$  at the point  $p$  may be thought of geometrically as a “tangent hyperplane” to  $N$  at  $p$ . We may construct a basis for the tangent space according to the following lemma:

**Lemma 2.1.1** *Let  $N$  be a smooth manifold of dimension  $n$ . Let  $p$  be any point of  $N$ . The tangent space  $T_p N$  to  $N$  at  $p$  is an  $n$ -dimensional vector space over the field  $\mathcal{R}$ . If  $(\phi_1, \dots, \phi_n)$  is a set of coordinates for  $N$  around  $p$ , then the tangent vectors  $(\frac{\partial}{\partial \phi_1})_p, \dots, (\frac{\partial}{\partial \phi_n})_p$  form a basis of  $T_p N$ , called the **natural basis** induced by the coordinates  $(\phi_1, \dots, \phi_n)$ .*

Lemma 2.1.1 provides a means of interpreting  $v(\lambda)$ , where  $v$  is a tangent vector at the point  $p$  and  $\lambda$  is a smooth function in the neighborhood of  $p$ . Let  $N = \mathcal{R}^n$  and let  $p \in N$ . If we choose the  $\phi_i$  in Lemma 2.1.1 to be the standard coordinate set in  $\mathcal{R}^n$ , i.e.,  $\phi_1 = [x_1 \ 0 \dots 0]^T, \dots, \phi_n = [0 \dots 0 \ x_n]^T$ , then from the definition of a basis we can write  $v(\lambda) = \sum_{i=1}^n v_i (\frac{\partial}{\partial x_i})_p(\lambda) = \sum_{i=1}^n (\frac{\partial \lambda}{\partial x_i})_p v_i$ , so that  $v(\lambda)$  can be seen to be the value of the derivative of  $\lambda$  along the direction of the vector  $[v_1 \dots v_n]^T$  at the point  $p$ .

Before defining a center manifold, we need a few facts from linear algebra. Suppose the Jacobian matrix  $J_0$  of (2.13) has  $n^c$  ( $c$  for center manifold) eigenvalues with zero real part,  $n^s$  ( $s$  for stable manifold) eigenvalues with negative real part, and  $n^u$  ( $u$  for unstable manifold) eigenvalues with positive real part. It is well known from linear algebra, e.g., see [28], that the domain of the linear mapping  $F$  can be decomposed into the direct sum of three invariant subspaces, noted  $E^c$ ,  $E^s$ , and  $E^u$  (with respective dimensions  $n^c$ ,  $n^s$ , and  $n^u$ ), with the property that  $J_0|_{E^c}$  ( $J_0$  with its domain restricted to  $E^c$ ) has all eigenvalues with zero real part,  $J_0|_{E^s}$  has all eigenvalues with negative real part, and  $J_0|_{E^u}$  has all eigenvalues with positive real part. Furthermore, it is true that the direct sum of these three subspaces equals the tangent space to  $N$  at  $x = 0$ ,  $T_0 N$ . We now have the machinery in place to define a center manifold.

Let  $x = 0$  be an equilibrium of (2.13). A manifold  $S$ , passing through  $x = 0$ , is said to be a **center manifold** for (2.13) at  $x = 0$ , if it is locally invariant and the tangent space to  $S$  at zero is exactly  $E^c$ .

In the remainder of this section, we will be interested in using Center Manifold Theory to determine if an equilibrium is stable or not. By the *Principle of Stability in the First Approximation*, we can declare an equilibrium unstable if  $J_0$  has any eigenvalues with positive real part. Thus, we will assume that  $J_0$  has eigenvalues with only zero or negative real parts. Under this assumption, we can write (2.13) in the form

$$\dot{y} = Ay + g(y, z) \quad (2.21)$$

$$\dot{z} = Bz + h(y, z) \quad (2.22)$$

where the eigenvalues of the matrix  $A$  all have negative real part and the eigenvalues of the matrix  $B$  all have real part equal to zero, by performing the following actions. Write  $a(x) = J_0x + \bar{a}(x)$ , and then reduce  $J_0$  to a block diagonal form

$$T^{-1}J_0T = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \quad (2.23)$$

by letting the columns of the matrix  $T$  be the right eigenvectors of  $J_0$  and defining the linear coordinate transformation

$$\begin{pmatrix} y \\ z \end{pmatrix} = T^{-1}x \quad (2.24)$$

The existence of center manifolds for (2.21), (2.22) is guaranteed by the following theorem (recall in (2.13) that  $a \in C^r, r \geq 2$ ).

**Theorem 2.1.7** *There exist a neighborhood  $U \subseteq \mathcal{R}^{n^c}$  of  $z = 0$  and a  $C^{r-1}$  mapping  $\pi : U \rightarrow \mathcal{R}^{n^s}$  such that*

$$S = \{(y, z) \in \mathcal{R}^{n^s} \times U \mid y = \pi(z)\} \quad (2.25)$$

*is a center manifold for (2.21), (2.22).*

**Proof:** See [10]. ■

Although existence of a center manifold for (2.21), (2.22) is guaranteed by Theorem 2.1.7, uniqueness is not, and in general, many center manifolds for a system are possible. Nevertheless, the above

definition and Theorem 2.1.7 are sufficient to characterize the critical behavior of any center manifold for (2.21), (2.22). By definition, a center manifold passes through  $(0, 0)$  and is tangent to  $E^c$  at  $x = 0$ . The tangency condition is equivalent to the mapping  $\pi$  having zero slope in the directions of the  $z$  coordinates. Thus,  $\pi$  must satisfy

$$\pi(0) = 0, \quad \frac{\partial \pi}{\partial z}(0) = 0 \quad (2.26)$$

The locally invariant nature of  $S$  places another constraint on the mapping  $\pi$ . Since  $y(t) = \pi(z(t))$  everywhere on  $S$ , we can differentiate this expression and use (2.21), (2.22) to obtain

$$\frac{dy}{dt} = A\pi(z(t)) + g(\pi(z(t)), z(t)) = \frac{\partial \pi}{\partial z} \frac{dz}{dt} = \frac{\partial \pi}{\partial z} (Bz(t) + h(\pi(z(t)), z(t)))$$

Thus,  $\pi$  must satisfy the partial differential equation

$$\frac{\partial \pi}{\partial z} (Bz + h(\pi(z), z)) = A\pi(z) + g(\pi(z), z) \quad (2.27)$$

Characterization of the system's behavior on the center manifold is critical because it can be shown that all system trajectories which begin in a neighborhood of the center manifold will approach it with exponential convergence [33]. Thus, system behavior on the center manifold determines local stability of the equilibrium point in question. This behavior can be determined by solving (2.22) with  $y$  set equal to  $\pi(z)$ , that is, solving

$$\dot{z} = Bz + h(\pi(z), z) \quad (2.28)$$

This is formalized in the following theorem.

**Theorem 2.1.8 (Reduction Principle)** *Suppose  $z = 0$  is a stable (respectively asymptotically stable, unstable) equilibrium of (2.28). Then  $(y, z) = (0, 0)$  is a stable (respectively asymptotically stable, unstable) equilibrium of (2.21), (2.22).*

**Proof:** See [10]. ■

This theorem is called the Reduction Principle because it reduces the dimension of the system that must be studied to determine stability of (2.21), (2.22) from  $n$  to  $n^c$ , which can be a major simplification. However, (2.28) must still be solved or its stability determined in some way. The Lyapunov theorems of Section 2.1.1 provide one tool for doing so. Alternatively, we introduce several other helpful theorems and lemmas in the next section.

### 2.1.3 Useful Theorems and Lemmas

Application of Lyapunov and center manifold theory to systems of various assumed general structures allow characterization of stability for these systems. We present several of these well known results here [33], to aid in later stability analyses.

**Lemma 2.1.2** *Let  $y = \pi_k(z)$  be a polynomial of degree  $k$ ,  $1 < k < r$ , satisfying*

$$\pi_k(0) = 0, \quad \frac{\partial \pi_k}{\partial z}(0) = 0$$

*and suppose*

$$\frac{\partial \pi_k}{\partial z}(Bz + h(\pi_k(z), z)) - A\pi_k(z) - g(\pi_k(z), z) = R_k(z)$$

*where  $R_k$  is some (possibly unknown) function vanishing at 0 together with all partial derivatives of order less than or equal to  $k$ . Then, any solution  $\pi(z)$  of (2.27) is such that the difference*

$$D_k(z) = \pi(z) - \pi_k(z)$$

*vanishes at 0 together with all partial derivatives of order less than or equal to  $k$ .*

Lemma 2.1.2 allows us to use the next lemma in determining stability of center manifold equations.

**Lemma 2.1.3** *Consider the one-dimensional system*

$$\dot{x} = kx^m + Q_m(x)$$

with  $m \geq 2$ ,  $0 \neq k \in \mathcal{R}$  and  $Q_m(x)$  a function vanishing at zero together with all partial derivatives of order less than or equal to  $m$ . The point of equilibrium  $x = 0$  is asymptotically stable if  $m$  is odd and  $k < 0$ . The equilibrium is unstable if  $m$  is odd and  $k > 0$ , or if  $m$  is even.

We next present several stability analysis tools for systems of the form of (2.21), (2.22), with various assumptions on  $g$  and  $h$ .

**Theorem 2.1.9** *Consider a system*

$$\begin{aligned}\dot{y} &= Ay + g(y, z) \\ \dot{z} &= h(y, z)\end{aligned}\tag{2.29}$$

and suppose that  $g(0, z) = 0 \ \forall \ z$  near 0 and

$$\frac{\partial g}{\partial y}(0, 0) = 0$$

If  $\dot{z} = h(0, z)$  has an asymptotically stable equilibrium at  $z = 0$  and the eigenvalues of  $A$  all have negative real part, then the system (2.29) has an asymptotically stable equilibrium at  $(y, z) = (0, 0)$ .

**Proof:** See [33], pg. 512. ■

**Lemma 2.1.4** *Consider a system*

$$\begin{aligned}\dot{y} &= g(y) \\ \dot{z} &= h(y, z)\end{aligned}\tag{2.30}$$

and suppose that  $\dot{y} = g(y)$  has an asymptotically stable equilibrium at  $y = 0$ . If  $\dot{z} = h(0, z)$  has an asymptotically stable equilibrium at  $z = 0$ , then the system (2.30) has an asymptotically stable equilibrium at  $(y, z) = (0, 0)$ .

**Lemma 2.1.5** *Consider a system*

$$\begin{aligned}\dot{y} &= g(y) \\ \dot{z} &= h(y, z)\end{aligned}\tag{2.31}$$

and suppose that  $(y, z) = (0, 0)$  is an equilibrium of (2.31),  $\dot{y} = g(y)$  has a stable equilibrium at  $y = 0$ , and  $\dot{z} = h(0, z)$  has an asymptotically stable equilibrium at  $z = 0$ . Then the equilibrium  $(y, z) = (0, 0)$  of (2.31) is stable.

We conclude this section by noting that, in all situations where an asymptotically stable equilibrium is required, the equilibrium does not necessarily have to be stable *in the first approximation*, i.e., have a Jacobian with all eigenvalues in the open complex left-half plane. Thus, center manifold or Lyapunov theory may be required to determine the asymptotic stability of an equilibrium, prior to invoking the above theorems and lemmas.

## 2.2 Feedback Linearization

This section provides a summary of the theory needed to conduct nonlinear control design via the technique of feedback linearization, also known as dynamic inversion. In this section we will consider input-affine multivariable nonlinear control systems in state space form, i.e.,

$$\begin{aligned}\dot{x} &= a(x) + \sum_{i=1}^m b_i(x)u_i \\ y &= c(x)\end{aligned}\tag{2.32}$$

where  $x \in U \subset \mathcal{R}^n$ ,  $u \in \mathcal{R}^m$ , and  $y \in \mathcal{R}^p$ . The mappings  $a, b_1, \dots, b_m$  are smooth,  $\mathcal{R}^n$ -valued mappings defined on the open set  $U$ , and may be represented by  $n$ -dimensional vectors of real-valued functions of the real variables  $x_1, \dots, x_n$ . A similar statement holds for  $c$  except  $p$  replaces  $n$ . Thus, we can write

$$a(x) = \begin{pmatrix} a_1(x_1, \dots, x_n) \\ \vdots \\ a_n(x_1, \dots, x_n) \end{pmatrix}, \quad b_i(x) = \begin{pmatrix} b_{i1}(x_1, \dots, x_n) \\ \vdots \\ b_{in}(x_1, \dots, x_n) \end{pmatrix}, \quad c(x) = \begin{pmatrix} c_1(x_1, \dots, x_n) \\ \vdots \\ c_p(x_1, \dots, x_n) \end{pmatrix}\tag{2.33}$$

We begin this section with definitions of needed terms. We then give the basic theory for single-input, single-output (SISO) systems in Section 2.2.2, and extend the results to the multi-input, multi-output (MIMO) case in Section 2.2.3. The bulk of the material is extracted from [33].

### 2.2.1 Terms and Definitions from Differential Geometry

Since  $a$  and  $b_i$  ( $c_i$ ) in (2.32) above map each point  $x$  of  $U$  to a vector in  $\mathcal{R}^n$  ( $\mathcal{R}^p$ ), they are often called **vector fields** defined on  $U$ . Similarly, a **covector field** defined on  $U$  is identified as a  $1 \times n$  (i.e. row) vector of functions of  $x$  (called a covector due to its association with the dual space of  $\mathcal{R}^n$ ). A special covector of importance is the **differential**, or **gradient**, of a real-valued function  $\lambda$  (also assumed to be defined on  $U$ ), which is defined in the standard way

$$d\lambda(x) = \frac{\partial \lambda}{\partial x} = \left[ \frac{\partial \lambda}{\partial x_1} \cdots \frac{\partial \lambda}{\partial x_n} \right] \quad (2.34)$$

We now define three types of differential operations involving smooth functions, vector fields, and covector fields used frequently in the analysis of nonlinear control systems and central to feedback linearization theory. Let  $\lambda$  and  $a$  be defined as above. Then the **Lie derivative** of  $\lambda$  along  $a$  (or with respect to  $a$ ), written  $L_a \lambda$ , is the smooth real-valued function defined by

$$L_a \lambda(x) = \langle d\lambda^T(x), a(x) \rangle = \frac{\partial \lambda}{\partial x} a(x) = \sum_{i=1}^n \frac{\partial \lambda}{\partial x_i} a_i(x) \quad (2.35)$$

for each  $x \in U$ , where  $\langle \cdot, \cdot \rangle$  denotes the standard inner or dot product on  $\mathcal{R}^n$ , and superscript “T” denotes the transpose operator. Note that this convention could have been used in all the theorems of Section 2.1.1, by replacing  $\dot{V}$  with  $L_a V$ . Also note that repeated use of this operation is possible, so that for example

$$L_g L_a \lambda(x) = \frac{\partial L_a \lambda}{\partial x} g(x)$$

If  $\lambda$  is being differentiated  $k$  times along  $a$ , the notation  $L_a^k \lambda$  is used, with  $L_a^0 \lambda = \lambda$ .

The second type of operation involves two vector fields  $a$  and  $g$ , both defined on an open subset  $U$  of  $\mathcal{R}^n$ . From these we construct a new smooth vector field called the **Lie bracket** or **Lie product** of  $a$  and  $g$ , denoted  $[a, g]$  (or  $ad_a g$ ), and defined at each  $x \in U$  as

$$[a, g] \equiv \frac{\partial g}{\partial x} a(x) - \frac{\partial a}{\partial x} g(x) \quad (2.36)$$

In (2.36) above,  $\partial a / \partial x$  and  $\partial g / \partial x$  represent the Jacobian matrices of  $a$  and  $g$ , respectively. Repeated bracketing of  $g$  with  $a$  is possible and can be defined recursively by

$$ad_a^k g(x) = [a, ad_a^{k-1} g](x)$$

for any  $k \geq 1$ , setting  $ad_a^0 g(x) = g(x)$ .

If  $a_i, g_i$ , and  $p$  are vector fields and  $r_i$  real numbers, it is easily proven that the Lie bracket operation possesses the following three properties

i. (Bilinearity)

$$[r_1 a_1 + r_2 a_2, g_1] = r_1 [a_1, g_1] + r_2 [a_2, g_1] \quad (2.37)$$

$$[a_1, r_1 g_1 + r_2 g_2] = r_1 [a_1, g_1] + r_2 [a_1, g_2] \quad (2.38)$$

ii. (Skew-Commutativity)

$$[a, g] = -[g, a] \quad (2.39)$$

iii. (Jacobi Identity)

$$[a, [g, p]] + [g, [p, a]] + [p, [a, g]] = 0 \quad (2.40)$$

The third type of operation involves a covector field  $w$  and a vector field  $v$  defined on an open subset  $U$  of  $\mathcal{R}^n$ . It produces a new covector field, called the **derivative of  $w$  along  $a$** , written  $L_a w$ , and defined for each  $x \in U$  as

$$L_a w(x) = a^T(x) \left( \frac{\partial w^T}{\partial x} \right)^T + w(x) \frac{\partial a}{\partial x} \quad (2.41)$$

Another important concept is that of a **distribution**. Suppose we are given  $l$  smooth vector fields  $a_1, \dots, a_l$ , with  $p$  elements each, all defined on an open subset  $U$  of  $\mathcal{R}^n$ . If we evaluate each  $a_i$  at some  $x \in U$ , we obtain  $l$  vectors, the span of which forms a subspace of  $\mathcal{R}^p$ . This smooth assignment of a point in  $U$  to a subspace of  $\mathcal{R}^p$  defines the rule for evaluation of a distribution,  $\Delta$ , i.e.,

$$\Delta(x) = \text{span}\{a_1(x), \dots, a_l(x)\} \quad \forall x \in U \quad (2.42)$$

and we use the notation

$$\Delta = \text{span}\{a_1, \dots, a_l\}$$

to define  $\Delta$  itself. Addition, intersection, and containment operations and relationships are defined pointwise for distributions according to normal subspace rules. Also, we say  $a$  **belongs to**  $\Delta$ , and

write  $a \in \Delta$ , if  $a(x) \in \Delta(x) \forall x \in U$ . The dimension of a distribution at a point  $x \in U$ ,  $\dim(\Delta(x))$ , is the dimension of the subspace  $\Delta(x)$ . A distribution  $\Delta$ , defined on an open set  $U$ , is said to be **nonsingular** if there exists an integer  $d$  such that

$$\dim(\Delta(x)) = d \forall x \in U$$

A point  $x^0$  of  $U$  is said to be a **regular point** of a distribution  $\Delta$ , if there exists a neighborhood  $U^0$  of  $x^0$  with the property that  $\Delta$  is nonsingular on  $U^0$ . A point of  $U$  which is not a regular point is called a **point of singularity**.

The following lemma illustrates the utility of a regular point of a distribution.

**Lemma 2.2.1** *Let  $\Delta$  be a smooth distribution and  $x^0$  a regular point of  $\Delta$ . Suppose  $\dim(\Delta(x^0)) = d$ . Then,  $\exists$  a neighborhood  $U^0$  of  $x^0$  and a set  $a_1, \dots, a_d$  of smooth vector fields defined on  $U^0$  such that*

*i. the vectors  $a_1(x), \dots, a_d(x)$  are linearly independent at each  $x \in U^0$*

*ii.  $\Delta(x) = \text{span}\{a_1(x), \dots, a_d(x)\} \forall x \in U^0$*

*iii. every smooth vector field  $\tau \in \Delta$  can be expressed on  $U^0$  as*

$$\tau(x) = \sum_{i=1}^d c_i(x) a_i(x)$$

*where  $c_1(x), \dots, c_d(x)$  are smooth real-valued functions of  $x$ , defined on  $U^0$ .*

A distribution  $\Delta$  is said to be **involutive** if the Lie bracket  $[\tau_1, \tau_2]$  of any pair of vector fields  $\tau_1$  and  $\tau_2$  belonging to  $\Delta$  is a vector field which belongs to  $\Delta$ , i.e., if

$$\tau_1 \in \Delta, \tau_2 \in \Delta \Rightarrow [\tau_1, \tau_2] \in \Delta$$

Thus, if  $\Delta$  is nonsingular,  $\Delta$  is involutive iff

$$[a_i, a_j] \in \Delta \forall i \geq 1, j \leq d \quad (2.43)$$

Because of (2.43), checking whether or not a nonsingular distribution is involutive amounts to checking if

$$\text{rank}(a_1(x) \dots a_d(x)) = \text{rank}(a_1(x) \dots a_d(x) [a_i, a_j]) \forall x, \forall i \geq 1, j \leq d$$

We note that a simple consequence of the above definitions is that any 1-dimensional distribution is involutive.

All of the above terms and definitions can similarly be posed in terms of covectors, thus resulting in **codistributions**. A useful example is the **annihilator** of the distribution  $\Delta(x)$ , defined by

$$\Delta^\perp(x) = \{w^* \in (\mathcal{R}^n)^* : \langle w^*, v \rangle = 0 \ \forall v \in \Delta(x)\}$$

and for which the following lemma applies.

**Lemma 2.2.2** *Let  $\Delta$  be a smooth distribution and  $x^0$  a regular point of  $\Delta$ . Then  $x^0$  is a regular point of  $\Delta^\perp$  and  $\exists$  a neighborhood  $U^0$  of  $x^0$  such that the restriction of  $\Delta^\perp$  to  $U^0$  is a smooth codistribution.*

A perhaps somewhat expected fact is that  $\dim(\Delta(x)) + \dim(\Delta^\perp(x)) = n$ .

As we have now built most of the necessary vocabulary, in the next section we proceed to introduce the basic concepts and theorems involved in feedback linearization.

## 2.2.2 Theory for Single-Input Single-Output Systems

In this section we present the basic theory of feedback linearization for SISO nonlinear dynamic systems. As mentioned in Section 1.1, the basic idea is to introduce a nonlinear change of coordinates and static state feedback which renders the closed loop system linear and stabilizable, enabling completion of the control design by application of any suitable linear design technique to the resulting system. Thus, we introduce some formal concepts about nonlinear coordinate changes in Section 2.2.2.1. Following that, we address the two cases in which at least some degree of linearization is possible: exact (total) linearization of the input to state response, and partial (input-output) linearization resulting in the so-called zero dynamics.

### 2.2.2.1 Local Coordinate Transformations

A global nonlinear change of coordinates can be described in the form

$$z = \Phi(x)$$

where  $\Phi(x)$  represents an  $\mathcal{R}^n$ -valued function of  $n$  variables, i.e.,

$$\Phi(x) = \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_n(x) \end{pmatrix} = \begin{pmatrix} \phi_1(x_1, \dots, x_n) \\ \vdots \\ \phi_n(x_1, \dots, x_n) \end{pmatrix}$$

with the following properties

- i.  $\Phi(x)$  is invertible, i.e. there exists a function  $\Phi^{-1}(z)$  such that

$$\Phi^{-1}(\Phi(x)) = x \quad \forall x \in \mathcal{R}^n \quad (2.44)$$

- ii.  $\Phi(x)$  and  $\Phi^{-1}(z)$  are both smooth mappings (have continuous partial derivatives of any order)

Such a global transformation is called a **global diffeomorphism** on  $\mathcal{R}^n$ . If (2.44) above only holds on some neighborhood of a point of interest, the transformation  $\Phi$  is called a **local diffeomorphism**.

The following theorem provides a means of checking whether or not a given mapping is a local diffeomorphism.

**Theorem 2.2.1 (Inverse Function)** *Suppose  $\Phi(x)$  is a smooth function defined on some subset  $U$  of  $\mathcal{R}^n$ . Suppose the Jacobian matrix of  $\Phi$  is nonsingular at a point  $x^0 \in U$ . Then, on a suitable open subset  $U^0$  of  $U$  containing  $x^0$ ,  $\Phi(x)$  defines a local diffeomorphism.*

**Proof:** See [1], Theorem 13.6. ■

We now focus our attention on obtaining conditions for existence of a special type of diffeomorphism to achieve our goal of feedback linearization. Prior to proceeding, we need one more definition. Recall Equation (2.32), where now  $m = p = 1$ , corresponding to our assumption of a SISO nonlinear dynamic system. Such a system is said to have **relative degree  $r$**  at a point  $x^0$  if

- i.  $L_b L_a^k c(x) = 0 \quad \forall x$  in a neighborhood of  $x^0$  and  $\forall k < r - 1$
- ii.  $L_b L_a^{r-1} c(x^0) \neq 0$

We note that there may exist points where a relative degree cannot be defined. However, the set of points where it is defined is an open and dense subset of the set  $U$  for which (2.32) is defined [33].

Intuitively speaking, the relative degree of a SISO system is the number of times the output  $y$  must be differentiated for the input to appear explicitly on the right hand side of the resulting expression. This can easily be seen from simple construction, i.e., given  $x^0 = x(t^0)$  we have

$$\begin{aligned} y(t^0) &= c(x(t^0)) = c(x^0) \\ y^{(1)}(t) &= \frac{\partial c}{\partial x} \frac{dx}{dt} = \frac{\partial c}{\partial x} (a(x(t)) + b(x(t))u(t)) \\ &= L_a c(x(t)) + L_b c(x(t))u(t) \end{aligned} \quad (2.45)$$

so that

$$\begin{aligned} y^{(k)}(t) &= L_a^k c(x(t)) \quad \forall k < r, \forall t \text{ near } t^0 \\ y^{(r)}(t^0) &= L_a^r c(x^0) + L_b L_a^{r-1} c(x^0)u(t^0) \end{aligned} \quad (2.46)$$

We also observe that if

$$L_b L_a^k c(x) = 0 \quad \forall k \geq 0, \forall x \text{ in a neighborhood of } x^0$$

then the output is not affected by the system input for all  $t$  near  $t^0$ , and no relative degree can be defined at any point around  $x^0$ . The concept of relative degree plays a central role in determining when we may find a local diffeomorphism *such that the output  $y = c(x)$  is a state variable in the new set of coordinates*, as indicated by the following lemma [33].

**Lemma 2.2.3** *Suppose the system (2.32) with  $m = p = 1$  has relative degree  $r$  at  $x^0$ . Then  $r \leq n$ .*

*Set*

$$\begin{aligned} \phi_1(x) &= c(x) \\ \phi_2(x) &= L_a c(x) \\ &\vdots \\ \phi_r(x) &= L_a^{r-1} c(x) \end{aligned} \quad (2.47)$$

If  $r < n$ ,  $\exists$   $n - r$  more functions  $\phi_{r+1}(x), \dots, \phi_n(x)$  such that the mapping

$$\Phi(x) = \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_n(x) \end{pmatrix}$$

has a nonsingular Jacobian matrix at  $x^0$ . The value of these additional functions at  $x^0$  can be fixed arbitrarily. Moreover, it is always possible to choose  $\phi_{r+1}(x), \dots, \phi_n(x)$  such that

$$L_b \phi_i(x) = 0 \quad \forall r+1 \leq i \leq n, \forall x \text{ near } x^0 \quad (2.48)$$

From Theorem 2.2.1, we see that  $\Phi$  in Lemma 2.2.3 qualifies as a local diffeomorphism. In concluding this section, we can say something more about the form of the system equations in the new set of coordinates obtained under the mapping  $\Phi$ . If (2.48) is satisfied, then the resulting state-space system description in the new set of coordinates is of the form

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_{r-1} &= z_r \\ \dot{z}_r &= \beta(z) + \alpha(z)u \\ \dot{z}_{r+1} &= q_{r+1}(z) \\ &\vdots \\ \dot{z}_n &= q_n(z) \end{aligned} \quad (2.49)$$

and we see that the first  $r - 1$  state equations are a chain of integrators, the  $r$ th state equation has the control entering directly, and the last  $n - r$  state equations show no direct effect of the control. A set of state equations in the form of (2.49) above is said to be in **normal form**. Lemma 2.2.3 states that it is always possible to satisfy (2.48), but in fact this may not always be easy. As per Theorem 2.2.1, if all we desire is a local diffeomorphism, we may in fact choose  $\phi_{r+1}(x), \dots, \phi_n(x)$

arbitrarily so long as the Jacobian of  $\Phi$  at  $x^0$  is nonsingular. If we do so, however, the resulting state equations will in general have the form

$$\begin{aligned}
 \dot{z}_1 &= z_2 \\
 \dot{z}_2 &= z_3 \\
 &\vdots \\
 \dot{z}_{r-1} &= z_r \\
 \dot{z}_r &= \beta(z) + \alpha(z)u \\
 \dot{z}_{r+1} &= q_{r+1}(z) + p_{r+1}(z)u \\
 &\vdots \\
 \dot{z}_n &= q_n(z) + p_n(z)u
 \end{aligned} \tag{2.50}$$

where now  $u$  enters the last  $n-r$  state equations explicitly. The benefits of choosing  $\phi_{r+1}(x), \dots, \phi_n(x)$  according to (2.48) will be seen in the next section, as we proceed toward obtaining a linearized input-output relationship.

#### 2.2.2.2 Exact Linearization Via Feedback

If we examine (2.49) a little more closely, we see that for a SISO system with relative degree  $r$  in normal form, the subsystem associated with the first  $r-1$  transformed coordinates is linear, while the effect of the control shows up only in the  $r$ th equation. Thus, by defining the static state feedback

$$u = \frac{1}{\alpha(z)}(-\beta(z) + v) \tag{2.51}$$

with  $v$  an arbitrary reference input, we end up with a closed loop system governed by the equations

$$\begin{aligned}
 \dot{z}_1 &= z_2 \\
 \dot{z}_2 &= z_3 \\
 &\vdots \\
 \dot{z}_{r-1} &= z_r
 \end{aligned}$$

$$\begin{aligned}
\dot{z}_r &= v \\
\dot{z}_{r+1} &= q_{r+1}(z) \\
&\vdots \\
\dot{z}_n &= q_n(z)
\end{aligned} \tag{2.52}$$

and now we see that the subsystem consisting of the first  $r$  equations has been rendered linear and controllable by the change of coordinates in Lemma 2.2.3 and application of the static state feedback defined in (2.51). Thus, if our system had relative degree  $r = n$ , by the use of the above transformation and state feedback, our entire set of dynamics would be rendered linear and controllable, accomplishing our stated purpose of feedback linearization. Achieving this complete linearization of all the state equations is known as solving the *SISO State Space Exact Linearization Problem*, and we have just shown that a sufficient condition for solving it is for our SISO system to have relative degree  $r = n$  at some point of interest. It turns out this is also a necessary condition, as formalized in the following theorem.

**Theorem 2.2.2** *The SISO State Space Exact Linearization Problem is solvable iff  $\exists$  a neighborhood  $U$  of  $x^0$  and a real-valued function  $\lambda(x)$ , defined on  $U$ , such that the SISO system*

$$\dot{x} = a(x) + b(x)u \tag{2.53}$$

$$y = \lambda(x) \tag{2.54}$$

*has relative degree  $n$  at  $x^0$ .*

**Proof:** See [33], Lemma 4.2.1. ■

As can be seen from Theorem 2.2.2, the solvability of the SISO State Space Exact Linearization Problem requires the existence of a suitable output function for which the system has relative degree  $n$ . It turns out that the existence of such a function can be tied to conditions on the vector fields  $a$  and  $b$  in (2.53). These conditions are specified in the following theorem.

**Theorem 2.2.3** Suppose a SISO system

$$\dot{x} = a(x) + b(x)u$$

is given. The SISO State Space Exact Linearization Problem is solvable near a point  $x^0$  (i.e.  $\exists$  an output function  $\lambda(x)$  for which the system has relative degree  $n$  at  $x^0$ ) iff the following conditions are satisfied:

- i. the matrix  $[b(x^0) \ ad_a b(x^0) \ \dots \ ad_a^{n-2} b(x^0) \ ad_a^{n-1} b(x^0)]$  has rank  $n$
- ii. the distribution  $D = \text{span}\{b, ad_a b, \dots, ad_a^{n-2} b\}$  is involutive near  $x^0$

**Proof:** See [33], Lemma 4.2.2. ■

Thus, given a SISO system (2.53), to solve the SISO State Space Exact Linearization Problem, we first check conditions (i) and (ii) of Theorem 2.2.3. If these are met, we next seek the output function  $\lambda(x)$ . It turns out that one can solve for  $\lambda$  by one of two methods, either straightforwardly from the definition of relative degree, i.e., solve

$$L_b \lambda(x) = L_b L_a \lambda(x) = \dots = L_b L_a^{n-2} \lambda(x) = 0 \ \forall \ x \text{ near } x^0 \quad (2.55)$$

$$L_b L_a^{n-1} \lambda(x^0) \neq 0 \quad (2.56)$$

or equivalently, solve

$$L_b \lambda(x) = L_{ad_a b} \lambda(x) = \dots = L_{ad_a^{n-2} b} \lambda(x) = 0 \ \forall \ x \text{ near } x^0 \quad (2.57)$$

$$L_{ad_a^{n-1} b} \lambda(x^0) \neq 0 \quad (2.58)$$

We then choose the coordinate transformation  $\Phi$  according to Lemma 2.2.3, i.e., set

$$\Phi(x) = [\lambda(x) \ L_a \lambda(x) \ \dots \ L_a^{n-1} \lambda(x)]^T$$

and construct the linearizing feedback

$$u = \phi(x) + \theta(x)v \quad (2.59)$$

by setting

$$\phi(x) = \frac{-L_a^n \lambda(x)}{L_b L_a^{n-1} \lambda(x)} \quad (2.60)$$

$$\theta(x) = \frac{1}{L_b L_a^{n-1} \lambda(x)} \quad (2.61)$$

We conclude this section by making a few remarks concerning the above procedure. First, the requirement of relative degree  $n$  assures that the denominators of (2.60) and (2.61) are nonzero, and thus that the linearizing feedback is well defined. Second, more than one appropriate output function  $\lambda$  may exist, as solutions to (2.55), (2.56), (2.57), and (2.58) are not unique. Third, we note that it can be shown that condition (i) of Theorem 2.2.3 is a type of controllability condition. In fact, it is equivalent to the requirement that the linear approximation of the system at  $x = 0$  be controllable [33].

The above procedure is the one to follow if the designer is free to choose the output function  $\lambda(x)$ . If given an output function  $c(x)$ , however, one simply determines the corresponding relative degree,  $r$ , of the system. If  $r$  is equal to  $n$ , a closed loop system with all  $n$  state equations linear and controllable is obtained, and the SISO State Space Exact Linearization Problem is solved. If  $r < n$ , then all of the state equations will not be rendered linear and controllable by the above procedure. In fact, we previously established that only  $r$  of the closed loop state equations will be rendered linear and controllable. We discuss the case where  $r < n$  in the next section.

### 2.2.2.3 Input-Output Linearization and Zero Dynamics

If we have a SISO system (2.53) and an output function  $\lambda(x)$  for which the relative degree  $r$  is strictly less than  $n$ , then Equation (2.52) shows that if we follow the procedure in the preceding section (namely coordinate transformation via Lemma 2.2.3 and state feedback via (2.59), (2.60), and (2.61) with  $r$  replacing  $n$ ), we obtain a closed loop system such that the first  $r$  equations will be linear and controllable, and the remaining  $n - r$  equations will be nonlinear and unaffected by the control. It can also be seen that the input-output behavior is totally captured by the first  $r$  state equations, and is therefore linear and controllable. We can then use any linear design technique we

choose to ensure that this  $r$ -dimensional subsystem of the state dynamics is stable and well behaved in terms of performance. Thus, the procedure enables us to obtain the desired linear and controllable relationship between input and output, but this solution is not complete because the last  $n - r$  state equations are uncontrollable and unobservable, and we must therefore consider their stability to ensure the overall system remains stable. We proceed to address this issue by introducing some simplifying notation and then invoking some previously established stabilization concepts.

We shall first decompose the transformed state vector into two groups: the first  $r$  state variables and the last  $n - r$  of them. Thus, let

$$\xi = \begin{bmatrix} z_1 \\ \vdots \\ z_r \end{bmatrix}, \quad \eta = \begin{bmatrix} z_{r+1} \\ \vdots \\ z_n \end{bmatrix}$$

Then the normal form of the above described system can be written more simply as

$$\dot{z}_1 = z_2$$

$$\dot{z}_2 = z_3$$

$$\vdots$$

$$\dot{z}_{r-1} = z_r$$

$$\dot{z}_r = \beta(\xi, \eta) + \alpha(\xi, \eta)u \quad (2.62)$$

$$\dot{\eta} = q(\xi, \eta) \quad (2.63)$$

Clearly, if we design the control to ensure stability of the linearized  $\xi$  subsystem (stability of  $\xi = 0$  in (2.62)), and if we can establish the local asymptotic stability of  $\eta = 0$  in (2.63), by methods such as those of Section 2.1, this will be sufficient to guarantee the local asymptotic stability of the overall closed loop system. It is not necessary, however, as the lemmas of Section 2.1.3 show. For, recalling Lemma 2.1.4, a sufficient condition for local asymptotic stability of the origin is that, if  $\xi = 0$  is LAS and if the **zero dynamics** of the system are LAS, i.e.,  $\eta = 0$  of

$$\dot{\eta} = q(0, \eta) \quad (2.64)$$

is LAS, then the origin  $(\xi, \eta) = (0, 0)$  of the overall system is LAS. The dynamics of (2.64) are called the zero dynamics because they describe the evolution of the state trajectory when the output  $y$  is constrained to equal zero for all time. It turns out that analyzing the stability of (2.64) is often much simpler than analyzing that of (2.63), and thus represents a significant simplification in required analysis. We conclude this section by remarking that the zero dynamics prove critical when the linear approximation of (2.64) at  $\eta = 0$  has no positive eigenvalues, else we could invoke the *Principle of Stability in the First Approximation* to conclude instability of the closed loop system.

#### 2.2.2.4 Other Interesting Results

In this section, we briefly summarize two other interesting problems and results involving feedback linearization of SISO systems: asymptotic output tracking and the disturbance decoupling problem. In the asymptotic output tracking problem, we are interested in producing a system output, which, irrespective of the initial system state, converges asymptotically to a prescribed reference output  $y_R(t)$ . It turns out that this may be achieved if a system has a well defined relative degree,  $r$ , in a neighborhood of  $y_R$ , and if certain other conditions are met. This can be seen from recalling the normal form equations, (2.62), (2.63), for a system with relative degree  $r$ , and choosing the control input

$$\begin{aligned} u &= \frac{1}{\alpha(\xi, \eta)} \left( -\beta(\xi, \eta) + y_R^{(r)} - \sum_{i=1}^r c_{i-1} (z_i - y_R^{(i-1)}) \right) \\ &= \frac{1}{L_b L_a^{r-1} c(x)} \left( -L_a^r c(x) + y_R^{(r)} - \sum_{i=1}^r c_{i-1} (L_a^{(i-1)} c(x) - y_R^{(i-1)}) \right) \end{aligned} \quad (2.65)$$

If we define  $e(t) = y(t) - y_R(t)$  and recall  $y = z_1$ , we obtain

$$\dot{z}_r = y^{(r)} = y_R^{(r)} - c_{r-1} e^{(r-1)} - \dots - c_1 e^{(1)} - c_0 e \quad (2.66)$$

or

$$e^{(r)} + c_{r-1} e^{(r-1)} + \dots + c_1 e^{(1)} + c_0 e = 0 \quad (2.67)$$

so that if the  $c_i$ 's are chosen to ensure that all the roots of the characteristic equation for (2.67) have negative real parts, the linearized portion of the normal form equations will exponentially converge

to the desired trajectory. We still must be concerned with the internal dynamics equations, however, as formalized in the following theorem.

**Theorem 2.2.4** *Suppose  $y_R(t), y_R^{(1)}(t), \dots, y_R^{(r-1)}(t)$  are defined for all  $t \geq 0$  and bounded, and let  $\xi_R(t) = [y_R(t) \ y_R^{(1)}(t) \ \dots \ y_R^{(r-1)}(t)]^T$ . Let  $\eta_R(t)$  denote the solution of*

$$\dot{\eta} = q(\xi_R(t), \eta) \quad (2.68)$$

*satisfying  $\eta_R(0) = 0$ . Suppose this solution is defined for all  $t \geq 0$ , bounded, and uniformly asymptotically stable, and the roots of the characteristic equation for (2.67) all have negative real parts. Then there exists a neighborhood of the total reference trajectory,  $[\xi_R^T(t) \ \eta_R^T(t)]^T$ , such that for all initial conditions located in this neighborhood, the closed loop system response will asymptotically converge to the reference trajectory.*

**Proof:** See [33], Proposition 4.5.1. ■

A similar but more elaborate procedure exists for achieving asymptotic model matching of a linear reference model. The reader is referred to [33] for details.

The disturbance decoupling problem involves a system of the form

$$\begin{aligned} \dot{x} &= a(x) + b(x)u + g(x)d \\ y &= c(x) \end{aligned} \quad (2.69)$$

where  $d$  is an undesired input, or disturbance, that we desire not to affect the system output  $y$  at all. Once again, it turns out that under certain conditions given below, this may be achieved using a static state feedback, and that the system having some relative degree  $r$  at a point of interest  $x^0$  provides one component to solving this problem, as formalized below.

**Theorem 2.2.5** *Suppose the SISO system (2.69) has some relative degree  $r$  at a point of interest  $x^0$ . The problem of finding a feedback  $u = \phi(x) + \theta(x)v$ , defined locally around  $x^0$ , with  $v$  an arbitrary reference input, such that the output of the system is decoupled from the disturbance can be solved iff*

$$L_g L_a^i c(x) = 0 \ \forall \ 0 \leq i \leq r-1, \forall \ x \text{ near } x^0 \quad (2.70)$$

If this is the case, one solution is

$$u = -\frac{L_a^r c(x)}{L_b L_a^{r-1} c(x)} + \frac{v}{L_b L_a^{r-1} c(x)}$$

**Proof:** See [33], Proposition 4.6.1. ■

### 2.2.3 Extensions to Multi-Input Multi-Output Systems

In this section we present an extension of the results of Section 2.2.2 to the full multivariable case [33]. Conceptually, very little changes. However, we must amend some of the SISO definitions to make sense in the MIMO setting. Following this, we present the multivariable generalizations of the SISO results previously shown.

Recall (2.32), defining a nonlinear dynamic MIMO system with  $m$  affine inputs and  $p = m$  outputs. Note that we have assumed a square system. Such a system is said to have (vector) relative degree  $\{r_1, \dots, r_m\}$  at a point  $x^0$  if

- i.  $\forall 1 \leq j \leq m, \forall k < r_i - 1, \forall 1 \leq i \leq m$ , and  $\forall x$  in a neighborhood of  $x^0$

$$L_{b_j} L_a^k c_i(x) = 0 \tag{2.71}$$

- ii. the  $m \times m$  matrix

$$T(x) = \begin{bmatrix} L_{b_1} L_a^{r_1-1} c_1(x) & \cdots & L_{b_m} L_a^{r_1-1} c_1(x) \\ L_{b_1} L_a^{r_2-1} c_2(x) & \cdots & L_{b_m} L_a^{r_2-1} c_2(x) \\ \vdots & \cdots & \vdots \\ L_{b_1} L_a^{r_m-1} c_m(x) & \cdots & L_{b_m} L_a^{r_m-1} c_m(x) \end{bmatrix} \tag{2.72}$$

is nonsingular at  $x = x^0$ .

We remark that this definition includes the definition given for a SISO system in Section 2.2.2.1, and that  $r_i$  is exactly the number of times one must differentiate the  $i$ th output at  $t = t^0$  in order to have at least one component of the input vector  $u(t^0)$  explicitly appearing in the right hand side of the expression. With this definition, Lemma 2.2.3 of Section 2.2.2.1 extends to the MIMO case in a straightforward manner as given below.

**Lemma 2.2.4** Suppose the system (2.32) with  $m = p$  has (vector) relative degree  $\{r_1, \dots, r_m\}$  at  $x^0$ . Then  $r = r_1 + \dots + r_m \leq n$ . Set, for  $1 \leq i \leq m$ ,

$$\begin{aligned}\phi_1^i(x) &= c_i(x) \\ \phi_2^i(x) &= L_a c_i(x) \\ &\vdots \\ \phi_{r_i}^i(x) &= L_a^{r_i-1} c_i(x)\end{aligned}\tag{2.73}$$

If  $r < n$ ,  $\exists$   $n - r$  more functions  $\phi_{r+1}(x), \dots, \phi_n(x)$  such that the mapping

$$\Phi(x) = [\phi_1^1(x) \dots \phi_{r_1}^1(x) \dots \phi_1^m(x) \dots \phi_{r_m}^m(x) \phi_{r+1}(x) \dots \phi_n(x)]^T$$

has a nonsingular Jacobian matrix at  $x^0$ . The value of these additional functions at  $x^0$  can be fixed arbitrarily. Moreover, if the distribution

$$D = \text{span}\{b_1, \dots, b_m\}$$

is involutive near  $x^0$  it is always possible to choose  $\phi_{r+1}(x), \dots, \phi_n(x)$  such that

$$L_{b_j} \phi_i(x) = 0 \quad \forall r+1 \leq i \leq n, \forall 1 \leq j \leq m, \forall x \text{ near } x^0\tag{2.74}$$

Again from Theorem 2.2.1, we see that  $\Phi$  in Lemma 2.2.4 qualifies as a local diffeomorphism. The only thing new in Lemma 2.2.4 is the requirement for the distribution  $D$  to be involutive in order to be able to satisfy (2.74). This is not really a new requirement, however, as it was mentioned previously that any 1-dimensional distribution is involutive, and so this requirement is automatically satisfied in the SISO case. We also note here that although we have assumed a square system, these results can be extended to nonsquare systems as long as condition ii in the definition of (vector) relative degree is replaced by the assumption that the matrix  $T(x)$  has full row rank at  $x^0$  [33].

With the coordinate tranformation defined above, we now revisit the concept of normal forms, explicitly characterizing those we will obtain in the MIMO case. Consistent with previous notation,

we first decompose the transformed state vector by letting

$$\xi^i = \begin{bmatrix} \xi_1^i \\ \vdots \\ \xi_{r_i}^i \end{bmatrix} = \begin{bmatrix} \phi_1^i(x) \\ \vdots \\ \phi_{r_i}^i(x) \end{bmatrix}$$

for  $1 \leq i \leq m$ , and set

$$\xi = \begin{bmatrix} \xi^1 \\ \vdots \\ \xi^m \end{bmatrix}$$

Also let

$$\eta = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_{n-r} \end{bmatrix} = \begin{bmatrix} \phi_{r+1}(x) \\ \vdots \\ \phi_n(x) \end{bmatrix}$$

and

$$\alpha_{ij}(\xi, \eta) = L_{b_j} L_a^{r_i-1} c_i(\Phi^{-1}(\xi, \eta)) \quad 1 \leq i, j \leq m$$

$$\beta_i(\xi, \eta) = L_a^{r_i} c_i(\Phi^{-1}(\xi, \eta)) \quad 1 \leq i \leq m$$

Then the normal form of the above described system for  $1 \leq i \leq m$  can be written more simply as

$$\begin{aligned} \dot{\xi}_1^i &= \xi_2^i \\ &\vdots \\ \dot{\xi}_{r_i-1}^i &= \xi_{r_i}^i \\ \dot{\xi}_{r_i}^i &= \beta_i(\xi, \eta) + \sum_{j=1}^m \alpha_{ij}(\xi, \eta) u_j \end{aligned} \quad (2.75)$$

If (2.74) is satisfied, then the final  $n - r$  state space equations in the new set of coordinates is of the form

$$\dot{\eta} = q(\xi, \eta) \quad (2.76)$$

while if (2.74) is not satisfied but we still have a valid coordinate transformation, we will have

$$\dot{\eta} = q(\xi, \eta) + p(\xi, \eta)u \quad (2.77)$$

With (2.75) in hand, we can now very easily give a sufficient condition for and a solution to the State Space Exact Linearization Problem for MIMO nonlinear systems of the form (2.32). If we assume (2.32) has a (vector) relative degree  $\{r_1, \dots, r_m\}$  at a point  $x^0$  and  $r = r_1 + \dots + r_m$  equals the dimension of the state space  $n$ , then since the matrix  $T(x)$  is nonsingular in a neighborhood of  $x^0$ , we can define the static state feedback

$$u = T^{-1}(\xi)[- \beta(\xi) + v] \quad (2.78)$$

with  $v$  an arbitrary  $m$ -vector to yield a closed loop system governed by the  $m$  sets of equations

$$\begin{aligned} \dot{\xi}_1^i &= \xi_2^i \\ &\vdots \\ \dot{\xi}_{r_i-1}^i &= \xi_{r_i}^i \\ \dot{\xi}_{r_i}^i &= v_i \end{aligned} \quad (2.79)$$

which are clearly linear and controllable. It turns out that the above condition on (vector) relative degree is also necessary, as in the SISO case, as long as the rank of  $b(x^0)$  is  $m$ . We now present the MIMO versions of Theorems 2.2.2 and 2.2.3.

**Theorem 2.2.6** *Suppose the matrix  $b(x^0)$  has rank  $m$ . Then, the MIMO State Space Exact Linearization Problem is solvable iff  $\exists$  a neighborhood  $U$  of  $x^0$  and  $m$  real-valued functions  $c_1(x), \dots, c_m(x)$ , defined on  $U$ , such that the system*

$$\dot{x} = a(x) + b(x)u \quad (2.80)$$

$$y = c(x) \quad (2.81)$$

*has some (vector) relative degree  $\{r_1, \dots, r_m\}$  at  $x^0$  and  $r_1 + r_2 + \dots + r_m = n$ .*

**Proof:** See [33], Lemma 5.2.1. ■

We remark that the rank  $m$  requirement for  $b(x^0)$  can be relaxed, and refer the reader to [33] for the details of how to accomplish this.

**Theorem 2.2.7** Suppose the matrix  $b(x^0)$  has rank  $m$ . Then,  $\exists$  a neighborhood  $U$  of  $x^0$  and  $m$  real-valued functions  $\lambda_1(x), \dots, \lambda_m(x)$ , defined on  $U$ , such that the system

$$\dot{x} = a(x) + b(x)u \quad (2.82)$$

$$y = \lambda(x) \quad (2.83)$$

has some (vector) relative degree  $\{r_1, \dots, r_m\}$  at  $x^0$ , with  $r_1 + r_2 + \dots + r_m = n$ , iff:

i. for each  $0 \leq i \leq n-1$ , the distribution  $D_i$  has constant dimension near  $x^0$ , where

$$D_i = \text{span}\{ad_a^k b_j : 0 \leq k \leq i, 1 \leq j \leq m\}$$

ii. the distribution  $D_{n-1}$  has dimension  $n$

iii. for each  $0 \leq i \leq n-2$ , the distribution  $D_i$  is involutive

**Proof:** See [33], Theorem 5.2.3. ■

Thus the MIMO State Space Exact Linearization Problem is solvable iff the three conditions in Theorem 2.2.7 are met. For a technique for constructing the appropriate output functions  $\lambda_i(x)$ , the reader is referred to [33].

Finally, as in Section 2.2.2, we need to consider the nonexact case, i.e., the case in which the input-output behavior may be linearized, but not all  $n$  of the transformed state equations are rendered linear. From our previous discussion of the MIMO normal form, (2.75), it is clear that if a system (2.32) possesses some vector relative degree at a point  $x^0$ , then the standard coordinate transformation and state feedback may be applied to yield an  $r = r_1 + \dots + r_m$ -dimensional subsystem that is linear and controllable, and if (2.74) is satisfied, an  $n-r$  dimensional subsystem unaffected by the control. As in Section 2.2.2.3, it is straightforward to apply Lemma 2.1.4 to conclude that if the zero dynamics of (2.75) are locally asymptotically stable, then the overall system may be rendered LAS by appropriate design of a stabilizing feedback control on the linearized portion of the state dynamics. It turns out that although having a vector relative degree is sufficient to obtain linearized input-output behavior in the MIMO case, it is not necessary as it was in the SISO case. In fact, a

broader class of MIMO systems may be input-output feedback linearized than just those possessing a vector relative degree. These systems are those for which Silverman's Structure Algorithm [60] may be successfully applied. This algorithm is rather complicated and will not be presented here, but the reader is referred to [33] for complete details.

In addition to those discussed for the SISO case, there are many other interesting results and problems unique to MIMO systems described by (2.32), only three of which will briefly be discussed here. It turns out that if a square ( $m \times m$ ) system (2.32) possesses a vector relative degree at a point  $x^0$ , this is sufficient to guarantee existence of a solution to the so-called *Noninteracting Control Problem* [33]. This problem is defined as achieving a closed loop system such that the existence of a vector relative degree is preserved and each control input affects one and only one separate output, i.e., the closed loop system is decomposed into  $m$  SISO systems. In fact, it is easy to see that the feedback control defined by (2.78) achieves the desired purpose.

From all the results presented thus far, it is clear that the existence of a vector relative degree for a system (2.32) is very useful for control design. When a given system does not possess such a relative degree, it may be possible to achieve one by augmenting the system with integrators on certain of the system inputs. This procedure is often called achieving relative degree via *Dynamic Extension*, and an algorithm may be found in [33]. A point of interest is that by this state variable augmentation, this procedure will result in a dynamic nonlinear controller, whereas up to this point, all other state feedback controllers considered have been static (nonlinear) feedbacks.

The final result we wish to present is not unique to the theory of feedback linearization, but instead is useful in all cases where we wish to achieve local asymptotic stabilization of a MIMO system such as (2.32). This result, taken from [33], relates stabilizability of the linear approximation of (2.32) at the origin

$$\dot{x} = Ax + Bu \tag{2.84}$$

where  $A$  equals the Jacobian of  $a$  at 0 and  $B = b(0)$ , to the stabilizability of the full nonlinear system, using either linear or nonlinear feedback.

**Theorem 2.2.8** *Suppose the linear approximation  $(A, B)$  of (2.84) is stabilizable. Then any linear feedback which asymptotically stabilizes the linear approximation also asymptotically stabilizes the original nonlinear system, at least locally. If the pair  $(A, B)$  is not controllable and there exist uncontrollable modes associated with eigenvalues with positive real part, then the original nonlinear system cannot be stabilized at all.*

**Proof:** See [33], Proposition 4.4.1. ■

Theorem 2.2.8 is easily proven by invoking the *Principle of Stability in the First Approximation*. The interesting aspect of the theorem is that stabilizability of the linearization is sufficient but not necessary for local asymptotic stabilization of a nonlinear system. Indeed, Theorem 2.2.8 implies that a nonlinear dynamic system whose linearization is not stabilizable may still be stabilizable in the nonlinear sense, provided any uncontrollable modes of the linearization have eigenvalues with zero real part. We will consider stabilization of just such a **critical system** in Chapter 3.

### 2.3 Recursive Backstepping

In this section we present the basic theory for nonlinear control design via the technique of recursive backstepping. Similar to Section 2.2, we start first with single-input systems, and then proceed to the multiple-input case. In fact, virtually all the existing results are for single-input systems. The great majority of the material is taken from [40].

#### 2.3.1 Theory for Single-Input Systems

In Section 2.1.1, we introduced the concept of a Lyapunov function for a system governed by the vector differential equation

$$\dot{x} = a(x, t)$$

The main idea was that the existence of a Lyapunov function was sufficient to guarantee at least the local stability of an equilibrium point of the system. Such a formulation is useful if one completely

knows the drift function  $a$ , and is simply seeking to determine system stability. If, however, we now consider the time-invariant system

$$\dot{x} = f(x, u), \quad x \in \mathcal{R}^n, \quad u \in \mathcal{R}, \quad f(0, 0) = 0 \quad (2.85)$$

where we are free to choose a feedback control input  $u = \alpha(x)$ , a natural question to ask is, does there exist a  $u$  such that the resulting closed loop system is globally asymptotically stable? If the answer is yes, the system (2.85) is said to possess a **control Lyapunov function** (CLF). Artstein [2] established a necessary and sufficient condition for the existence of a CLF for a system of the form of (2.85), presented below.

**Lemma 2.3.1** *A smooth, positive definite, and radially unbounded function  $V : \mathcal{R}^n \rightarrow \mathcal{R}_+$  is a control Lyapunov function for (2.85) if*

$$\inf_{u \in \mathcal{R}} \left\{ \frac{\partial V}{\partial x}(x) f(x, u) \right\} < 0 \quad \forall x \neq 0 \quad (2.86)$$

This intuitively simple concept of finding a feedback control and associated Lyapunov function for the resulting closed loop system equations forms the basis of control Lyapunov function design theory. The problem is that (2.86) speaks to the existence of the pair  $u$  and  $V$ , but provides little insight into how to choose them simultaneously to guarantee that the closed loop system is GAS. Recursive backstepping is one attempt to remedy this problem. It provides a systematic yet flexible procedure for sequentially building up the feedback control and CLF to accomplish the desired stability objective. Actually, recursive backstepping requires slightly different assumptions for a system than does control Lyapunov function theory. In fact, for an affine single-input system, it requires the following assumption.

**Assumption 2.3.1** *Consider the system*

$$\dot{x} = a(x) + b(x)u, \quad a(0) = 0 \quad (2.87)$$

where  $x \in \mathcal{R}^n$  is the state and  $u \in \mathcal{R}$  is the control input. There exist a continuously differentiable feedback law

$$u = \alpha(x), \quad \alpha(0) = 0 \quad (2.88)$$

and a smooth, positive definite, radially unbounded function  $V : \mathcal{R}^n \rightarrow \mathcal{R}_+$  such that

$$\frac{\partial V}{\partial x}(x)[a(x) + b(x)\alpha(x)] \leq -W(x) \leq 0, \forall x \in \mathcal{R}^n \quad (2.89)$$

where  $W : \mathcal{R}^n \rightarrow \mathcal{R}_+$  is positive semidefinite.

Under this assumption, the control (2.88), applied to the system (2.87), guarantees via the Lasalle-Yoshizawa theorem (Theorem 2.1.2) the global boundedness of  $x(t)$  and the asymptotic regulation of  $W(x(t))$ . Lasalle's Invariance theorem (Theorem 2.1.4) with  $\Omega = \mathcal{R}^n$  may also be applied to conclude that  $x(t)$  converges to the largest invariant set  $M$  contained in the set  $E = \{x \in \mathcal{R}^n \mid W(x) = 0\}$ . If we take  $W(x)$  in (2.89) to be positive definite instead of positive semidefinite (as is the assumption for CLF theory), then clearly the control (2.88) renders  $x = 0$  the GAS equilibrium of (2.87). Besides the different assumptions on the definiteness of  $W(x)$ , the other difference between recursive backstepping and CLF theory is that we have required the control (2.88) to be continuously differentiable in Assumption 2.3.1, whereas there is no such requirement in CLF theory. This  $C^1$  property of the control is crucial to the recursive backstepping design procedure, as illustrated in the following basic theorem.

**Theorem 2.3.1 (Integrator Backstepping)** *Let the system (2.87) be augmented by an integrator on the control input:*

$$\dot{x} = a(x) + b(x)\xi \quad (2.90)$$

$$\dot{\xi} = u \quad (2.91)$$

and suppose that (2.90) satisfies Assumption 2.3.1 with  $\xi \in \mathcal{R}$  as its control.

i. *If  $W(x)$  is positive definite, then*

$$V_a(x, \xi) = V(x) + \frac{1}{2}[\xi - \alpha(x)]^2 \quad (2.92)$$

*is a CLF for the full system (2.90), (2.91). Thus,  $\exists$  a feedback control  $u = \alpha_a(x, \xi)$  which renders  $(x, \xi) = (0, 0)$  the GAS equilibrium of the system. One such control is*

$$u = -c(\xi - \alpha(x)) + \frac{\partial \alpha}{\partial x}(x)[a(x) + b(x)\xi] - \frac{\partial V}{\partial x}(x)b(x), \quad c > 0 \quad (2.93)$$

- ii. If  $W(x)$  is only positive semidefinite, then  $\exists$  a feedback control which renders  $\dot{V}_a \leq -W_a(x, \xi) \leq 0$ , such that  $W_a(x, \xi) > 0$  whenever  $W(x) > 0$  or  $\xi \neq \alpha(x)$ . This guarantees global boundedness and convergence of  $[x^T(t) \ \xi(t)]^T$  to the largest invariant set  $M_a$  contained in the set  $E_a = \{[x^T \ \xi]^T \in \mathcal{R}^{n+1} \mid W(x) = 0, \ \xi = \alpha(x)\}$ .

The proof of Theorem 2.3.1 is very instructive of the procedure in general and allows us to define some necessary terms. We therefore include it in its entirety [40].

**Proof:** Examining Equation (2.90), we see that  $\xi$  now plays the role that  $u$  did in Assumption 2.3.1. We therefore consider  $\xi$  to be a **virtual control** for (2.90), and we choose a desired feedback control function  $\alpha(x)$  that will result in closed loop stability for (2.90) if we are able to make  $\xi = \alpha(x)$  by appropriately choosing  $u$  in (2.91). Accordingly,  $\alpha(x)$  is called a **stabilizing function**. We recognize, however, that we will never achieve perfect equality, and so we introduce the **error variable**

$$e = \xi - \alpha(x) \quad (2.94)$$

to aid in the analysis. Differentiating (2.94) with respect to time, we can rewrite the system equations as

$$\begin{aligned} \dot{x} &= a(x) + b(x)[\alpha(x) + e] \\ \dot{e} &= u - \frac{\partial \alpha}{\partial x}(x)[a(x) + b(x)(\alpha(x) + e)] \end{aligned} \quad (2.95)$$

Using (2.89), the derivative of (2.92) along the solutions of (2.95) is

$$\begin{aligned} \dot{V}_a &= \frac{\partial V}{\partial x}(a + b\alpha + be) + e \left[ u - \frac{\partial \alpha}{\partial x}(a + b(\alpha + e)) \right] \\ &= \frac{\partial V}{\partial x}(a + b\alpha) + e \left[ u - \frac{\partial \alpha}{\partial x}(a + b(\alpha + e)) + \frac{\partial V}{\partial x}b \right] \\ &\leq -W(x) + e \left[ u - \frac{\partial \alpha}{\partial x}(a + b(\alpha + e)) + \frac{\partial V}{\partial x}b \right] \end{aligned} \quad (2.96)$$

where the terms containing  $e$  have been grouped together. By the Lasalle-Yoshizawa theorem, any choice of the control  $u$  which renders  $\dot{V}_a \leq -W_a(x, \xi) \leq -W(x)$ , with  $W_a$  positive definite in  $e$ , guarantees global boundedness of  $x, e$ , and  $\xi$ , and regulation of  $W(x(t))$  and  $e(t)$ . Furthermore,

Lasalle's Invariance theorem guarantees convergence of  $[x^T(t) \ \xi(t)]^T$  to the largest invariant set contained in the set  $\{[x^T \ e]^T \in \mathcal{R}^{n+1} \mid W(x) = 0, e = 0\}$ . An obvious way to make  $\dot{V}_a$  negative definite in  $e$  is to choose the control as in (2.93), which renders the braced term in (2.96) equal to  $-ce$  and yields

$$\dot{V}_a \leq -W(x) - ce^2 \equiv -W_a(x, \xi) \leq -W(x) \leq 0 \quad (2.97)$$

Clearly, if  $W(x)$  is positive definite, Theorem 2.1.2 guarantees the global asymptotic stability of  $(x, e) = (0, 0)$ , which in turn implies that  $V_a(x, \xi)$  is a CLF and  $(x, \xi) = (0, 0)$  is the GAS equilibrium of (2.95). ■

While the choice of control (2.93) is simple, it is not unique and may not even be the best choice when considering other factors such as control usage, because it cancels all the nonlinearities in the braced term in (2.96), some of which may be beneficial to stability or dominated by more powerful terms, and thus not require cancellation. The authors of [40] state that the main result of backstepping is not a specific control law, but the construction of a Lyapunov function whose derivative can be made negative definite by a variety of control laws. The designer thus has significant freedom when using this method, and must exercise good engineering judgment when selecting the control.

Although Theorem 2.3.1 shows only one integrator, if the original control  $u$  in (2.87) is  $k$  times continuously differentiable, it is straightforward to extend the method of integrator backstepping to a chain of  $k$  augmented integrators, with the associated introduction of  $k$  stabilizing functions selected sequentially, starting from the original equation, and  $k$  error variables, quadratic positive definite terms of which are added to the original  $V(x)$  to construct the overall system Lyapunov function. While integrator backstepping can be quite useful for design, a more general technique may be needed. The next theorem provides the necessary generality.

**Theorem 2.3.2 (Nonlinear Block Backstepping)** *Consider the cascade system:*

$$\dot{x} = a(x) + b(x)y \quad (2.98)$$

$$\dot{\xi} = m(x, \xi) + \beta(x, \xi)u, \quad y = c(\xi), \quad c(0) = 0, \quad \xi \in \mathcal{R}^q, \quad u \in \mathcal{R} \quad (2.99)$$

Assume that (2.99) has globally defined and constant relative degree one uniformly in  $x$ , and that its zero dynamics subsystem is ISS with respect to  $x$  and  $y$  as inputs. Also suppose that (2.98) satisfies Assumption 2.3.1 with  $y$  as its control, except  $V(x)$  may be only positive semidefinite, and we require the closed loop system solution  $x(t)$  to be bounded if  $V(x(t))$  is bounded. Then there exists a feedback control which guarantees global boundedness and convergence of  $[x^T(t) \ \xi^T(t)]^T$  to the largest invariant set  $M_a$  contained in the set  $E_a = \{[x^T \ \xi^T]^T \in \mathcal{R}^{n+q} \mid W(x) = 0, y = \alpha(x)\}$ .

One such control is

$$u = \left( \frac{\partial c}{\partial \xi}(\xi) \beta(x, \xi) \right)^{-1} \cdot \left\{ -c_1(y - \alpha(x)) - \frac{\partial c}{\partial \xi}(\xi) m(x, \xi) + \frac{\partial \alpha}{\partial x}(x) [a(x) + b(x)y] - \frac{\partial V}{\partial x}(x) b(x) \right\}, \quad c_1 > 0 \quad (2.100)$$

Moreover, if both  $V(x)$  and  $W(x)$  are positive definite, then the equilibrium  $(x, \xi) = (0, 0)$  is GAS.

**Proof:** See [40], Lemma 2.25. ■

### 2.3.2 Theory for Multi-Input Systems

The entire recursive backstepping theory for multi-input systems to be found in [40] consists of a single design procedure for systems in what the authors call *multi-input parametric strict-feedback form*. This procedure includes provisions for parametric uncertainty in the system equations, which we will not cover here. In what follows, this author has therefore adapted the procedure to deal with nominal control design only, which applies for systems in the resulting *multi-input strict-feedback form*:

$$\begin{aligned} \dot{x}_1^i &= x_2^i \\ &\vdots \\ \dot{x}_{r_i-1}^i &= x_{r_i}^i \\ \dot{x}_{r_i}^i &= \sum_{j=1}^m \alpha_{ij}(x) u_j \end{aligned} \quad (2.101)$$

where  $1 \leq i \leq m$ ,  $\sum r_i = n$ , and we have the added condition that the control input matrix  $T$  is nonsingular for all  $x \in \mathcal{R}^n$ . Upon comparing (2.101) with (2.75), we see that apart from there

being no drift term  $\beta(x)$  in the last equation of (2.101), the two sets of equations are identical, as long as  $\sum r_i = n$  in (2.75) as well. Thus, the *multi-input strict-feedback form* can be obtained for any system possessing a (vector) relative degree  $\{r_1, \dots, r_m\}$  such that  $\sum r_i = n$ ; i.e., the system is exactly state space feedback linearizable. The multi-input strict-feedback recursive backstepping procedure for a system with  $m$  inputs is given in the following theorem.

**Theorem 2.3.3 (Multi-Input Strict-Feedback Recursive Backstepping)** *For the  $i$ th subsystem of (2.101),  $1 \leq i \leq m$ , define for  $1 \leq k \leq r_i$  the error variables,*

$$e_k^i = x_k^i - \alpha_{k-1}^i(x_1^i, \dots, x_{k-1}^i) \quad (2.102)$$

*and the stabilizing functions*

$$\alpha_k^i = -c_k^i e_k^i - e_{k-1}^i + \sum_{j=1}^{k-1} \frac{\partial \alpha_{k-1}^i}{\partial x_j^i}(x_{j+1}^i) \quad (2.103)$$

*where  $e_0^i \equiv \alpha_0^i \equiv 0$ , and  $c_k^i > 0$  are design constants. Also, let the control input be given by*

$$u = -T^{-1}(x) \begin{bmatrix} c_{r_1}^1 e_{r_1}^1 + e_{r_1-1}^1 \\ \vdots \\ c_{r_m}^m e_{r_m}^m + e_{r_m-1}^m \end{bmatrix} \quad (2.104)$$

*Then  $x(t)$  and  $e(t)$  are globally bounded and  $e(t) \rightarrow 0$  as  $t \rightarrow \infty$ . Furthermore,  $x = 0$  is a GAS equilibrium of the closed loop system.*

**Proof:** See [40], Theorem 3.5 and Section 3.3.2. ■

## 2.4 Nonlinear $H_\infty$ Control

In this section we present the modern theory of nonlinear  $H_\infty$  control design, perhaps more appropriately called nonlinear induced  $L_2$  control design. Our treatment begins with the problem definition and required assumptions, and then we briefly discuss some concepts related to the notion of induced  $L_2$  gain for nonlinear systems. Following that, we present sufficient conditions for achieving the desired control objective under both full state feedback and output feedback assumptions. The majority of this material can be found in [34] and [67].

### 2.4.1 Problem Definition

In this section we consider general nonlinear systems of the form

$$\begin{aligned}\dot{x} &= f(x, d, u) \\ y &= c(x, d) \\ z &= h(x, u)\end{aligned}\tag{2.105}$$

with state vector  $x \in \mathcal{R}^n$ , exogenous input  $d \in \mathcal{R}^r$ , control input  $u \in \mathcal{R}^m$ , measurement vector  $y \in \mathcal{R}^p$ , and penalized variable  $z \in \mathcal{R}^s$ . We note that  $d$  may contain both disturbances to be rejected (process and measurement noises) and reference signals to be tracked, and that  $z$  may contain both tracking type errors and penalties on control usage. We assume  $f, h$ , and  $c$  to be smooth vector valued functions defined in a neighborhood of the origin in  $\mathcal{R}^n \times \mathcal{R}^p \times \mathcal{R}^s$ . We also assume the existence of a fixed equilibrium for (2.105), and, as discussed in Section 2.1.1, we may without loss of generality take it to be the origin so that  $f(0, 0, 0) = 0$ ,  $h(0, 0) = 0$ , and  $c(0, 0) = 0$ . Now, let  $\gamma$  be a fixed positive constant. Then the nonlinear  $H_\infty$  *suboptimal control problem* (for disturbance attenuation level  $\gamma$ ) is to find a nonlinear (dynamic) compensator

$$\begin{aligned}\dot{\xi} &= k(\xi, y) \\ u &= m(\xi, y)\end{aligned}\tag{2.106}$$

with state vector  $\xi$ , and  $k$  and  $m$  smooth functions satisfying  $k(0, 0) = 0$ ,  $m(0, 0) = 0$ , such that the closed loop system resulting from (2.105) and (2.106) has induced  $L_2$  gain less than or equal to  $\gamma$ , where the definition of induced  $L_2$  gain for a nonlinear system is as follows. A system (2.105), (2.106), has induced  $L_2$  gain less than or equal to  $\gamma$  if for all initial conditions  $x(0), z(0)$ , there exists a nonnegative constant  $K$  (depending on  $x(0), z(0)$  and equal to zero for  $x(0) = 0, z(0) = 0$ ) such that

$$\int_0^T \|z(t)\|^2 dt \leq \gamma^2 \int_0^T \|d(t)\|^2 dt + K\tag{2.107}$$

for all  $d \in L_2[0, T]$  and  $T \geq 0$ , with  $z(t)$  denoting the response for initial condition  $x(0), \xi(0)$ . Then the nonlinear  $H_\infty$  *optimal control problem* is to find the smallest  $\gamma^* \geq 0$  such that the nonlinear  $H_\infty$

suboptimal control problem is solvable for all  $\gamma > \gamma^*$ . In the next section we provide a connection between the theory of dissipative systems and the induced  $L_2$  gain, which will allow us to derive desirable stability properties for systems with finite induced  $L_2$  gain.

### 2.4.2 Induced $L_2$ Gain Analysis and Dissipativity for Nonlinear Systems

Following standard convention [7, 27, 67], a nonlinear system

$$\begin{aligned}\dot{x} &= f(x, d) \\ z &= h(x)\end{aligned}\tag{2.108}$$

is said to be **locally dissipative** near  $(x, d) = (0, 0)$  with respect to the **supply rate**  $s(d, z)$  if there exists a positive semidefinite function  $V(x)$  such that for all  $x, u$  in a neighborhood,  $U$ , of  $(0, 0)$ , the *dissipation inequality*

$$V(x(T; x(0), 0)) \leq V(x(0)) + \int_0^T s(d(t), z(t)) dt \tag{2.109}$$

holds  $\forall T \geq 0$ .  $V$  above is called a **storage function** for (2.108). A system is said to be **dissipative** if (2.109) holds and  $U = \mathcal{R}^n \times \mathcal{R}^r$ . If we let the supply rate  $s = \gamma^2 \|d\|^2 - \|z\|^2$ , we see (recalling  $V$  is positive semidefinite) that (2.109) becomes

$$0 \leq V(x(T; x(0), 0)) \leq V(x(0)) + \int_0^T \gamma^2 \|d(t)\|^2 - \|z(t)\|^2 dt \tag{2.110}$$

which implies

$$\int_0^T \|z(t)\|^2 dt \leq \int_0^T \gamma^2 \|d(t)\|^2 dt + V(x(0)) \tag{2.111}$$

Thus, a system (2.108) which is (locally) dissipative with respect to the supply rate  $\gamma^2 \|d\|^2 - \|z\|^2$  has (locally) induced  $L_2$  gain less than or equal to  $\gamma$ .

If the storage function  $V$  is  $C^1$ , (2.109) is equivalent to the *differential dissipation inequality*

$$\frac{\partial V}{\partial x} f(x, d) - s(d, z) \leq 0 \tag{2.112}$$

or, again letting  $s = \gamma^2 \|d\|^2 - \|z\|^2$

$$\frac{\partial V}{\partial x} f(x, d) + \|z\|^2 - \gamma^2 \|d\|^2 \leq 0 \tag{2.113}$$

A consequence of (2.113) is that for  $d = 0$ , we get

$$\frac{\partial V}{\partial x} f(x, 0) \leq -\|z\|^2 \leq 0 \quad (2.114)$$

so that if the storage function  $V$  is positive definite, it serves as a Lyapunov function, and we obtain (local) asymptotic stability of the origin if the inequality in (2.114) is satisfied strictly. Alternatively, it turns out that  $V$  is indeed positive definite if (2.108) is assumed to be **zero-state observable**, i.e.,  $z(t) = 0 \ \forall t \geq 0 \Rightarrow x(0) = 0$  [67]. Moreover, if  $V$  is **proper**, i.e., for each  $c > 0$  the set  $\{x \in \mathcal{R}^n \mid 0 \leq V \leq c\}$  is compact, then the origin is GAS. If  $V$  is only positive semidefinite, then we still may be able to apply Theorem 2.1.4 to deduce desirable stability properties for (2.108), if the system meets certain requirements. One such case is that of a **zero-state detectable** system, which is one for which  $z(t) = 0 \ \forall t \geq 0 \Rightarrow x(t) \rightarrow 0$  as  $t \rightarrow \infty$  [34].

We close this section with a final comment relating the induced  $L_2$  gain of an input-affine non-linear system

$$\begin{aligned} \dot{x} &= a(x) + g(x)d, \quad a(0) = 0 \\ z &= h(x), \quad h(0) = 0 \end{aligned} \quad (2.115)$$

to that of its linearization at zero

$$\begin{aligned} \dot{x} &= Ax + Gd \\ z &= Hx \end{aligned} \quad (2.116)$$

where  $A = \frac{\partial a}{\partial x}(0)$ ,  $G = g(0)$ , and  $H = \frac{\partial h}{\partial x}(0)$ . Van der Schaft has shown [66] that if  $A$  is asymptotically stable and (2.116) has induced  $L_2$  gain  $< \gamma$ , then so does (2.115) locally, and vice versa. Now, the induced  $L_2$  gain of (2.116) is easily computable [56] from the bounded real lemma algebraic Riccati equation

$$A^T P + PA + \frac{1}{\gamma^2} PGG^T P + H^T H = 0 \quad (2.117)$$

by finding the smallest  $\gamma$  such that (2.117) has a symmetric, positive semidefinite solution  $P$  with the eigenvalues of  $A + \frac{1}{\gamma^2} GG^T P$  all having negative real parts. Thus, this result offers an easy way to

bound the local induced  $L_2$  gain of (2.115), and provides a check for any algorithm we may compose to estimate the optimal induced  $L_2$  gain of a nonlinear system.

### 2.4.3 State Feedback

In this section we consider the nonlinear  $H_\infty$  suboptimal control problem under the assumption of full state feedback. Thus, we seek a feedback law of the form  $u = \alpha(x)$  such that the closed loop system

$$\begin{aligned}\dot{x} &= f(x, d, \alpha(x)) \\ z &= h(x, \alpha(x))\end{aligned}\tag{2.118}$$

is (locally) dissipative with respect to the supply rate  $s(d, z) = \gamma^2 \|d\|^2 - \|z\|^2$ . This problem may be cast as a two-player, zero-sum, differential game, in which the minimizing player controls the input  $u$  and the maximizing player controls the exogenous input  $d$  [4]. We associate with this game a Hamiltonian function  $\mathcal{H} : \mathcal{R}^n \times \mathcal{R}^n \times \mathcal{R}^r \times \mathcal{R}^m \rightarrow \mathcal{R}$  defined as

$$\mathcal{H}(x, p, d, u) = p^T f(x, d, u) + \|h(x, u)\|^2 - \gamma^2 \|d\|^2\tag{2.119}$$

If we now assume the plant (2.118) satisfies the following hypothesis:

**Assumption 2.4.1** *The penalty map  $h(x, u)$  is such that the matrix*

$$D_1 = \frac{\partial h}{\partial u}(0, 0)$$

*has rank  $m$*

i.e., we place a nontrivial independent linear penalty on every control, then the Hessian matrix of  $\mathcal{H}$ , viewed as a function of  $(d, u)$ , at  $(x, p, d, u) = (0, 0, 0, 0)$  equals

$$\begin{bmatrix} -2\gamma^2 I & 0 \\ 0 & 2D_1^T D_1 \end{bmatrix}\tag{2.120}$$

Thus, the Hamiltonian  $\mathcal{H}(x, p, d, u)$  is quadratic in  $d$  and  $u$ , and since by Assumption 2.4.1,  $D_1^T D_1$  is positive definite, (2.120) implies  $\mathcal{H}$  has a unique local saddle point in  $(d, u)$  for each  $(x, p)$ . Thus, there

exist unique smooth functions  $d_*(x, p)$  and  $u_*(x, p)$ , defined in a neighborhood of  $(0, 0)$ , satisfying

$$\begin{aligned}\frac{\partial \mathcal{H}}{\partial d}(x, p, d_*, u_*) &= 0 \\ \frac{\partial \mathcal{H}}{\partial u}(x, p, d_*, u_*) &= 0 \\ d_*(0, 0) &= 0 \\ u_*(0, 0) &= 0\end{aligned}\tag{2.121}$$

and such that

$$\mathcal{H}(x, p, d, u_*(x, p)) \leq \mathcal{H}_*(x, p) \equiv \mathcal{H}(x, p, d_*(x, p), u_*(x, p)) \leq \mathcal{H}(x, p, d_*(x, p), u)\tag{2.122}$$

for each  $(x, p, d, u)$  in a neighborhood of  $(x, p, d, u) = (0, 0, 0, 0)$ .

If we now define  $V : \mathcal{R}^n \rightarrow \mathcal{R}_+$  to be a  $C^1$ , positive semidefinite function defined on a neighborhood  $U$  of  $x = 0$  and let  $V_x^T = \frac{\partial V}{\partial x}^T$  take the place of  $p$  in (2.122) and (2.121), we get

$$\mathcal{H}(x, V_x^T(x), d, \alpha_2(x)) \leq \mathcal{H}_*(x, V_x^T(x)) \leq \mathcal{H}(x, p, \alpha_1(x), u)\tag{2.123}$$

where we have defined

$$\alpha_1(x) \equiv d_*(x, V_x^T(x)), \quad \alpha_2(x) \equiv u_*(x, V_x^T(x))\tag{2.124}$$

Thus, if we satisfy the so-called *Hamilton-Jacobi Inequality*

$$\mathcal{H}_*(x, V_x^T(x)) \leq 0\tag{2.125}$$

for all  $x$  in a neighborhood of 0 and set  $u = \alpha_2(x)$  in (2.118), we obtain a closed loop system such that

$$V_x(x)f(x, d, \alpha_2(x)) + \|h(x, \alpha_2(x))\|^2 - \gamma^2 \|d\|^2 \leq 0\tag{2.126}$$

In solving (2.125) we have thus obtained a closed loop system with the required (local) dissipativity property, and as shown in Section 2.4.2, we have therefore solved the local nonlinear state feedback suboptimal  $H_\infty$  control problem. Closed loop stability may be guaranteed by imposing any of the conditions listed at the end of Section 2.4.2: zero-state detectability of (2.118) or  $V$  positive definite and either solution of (2.125) with strict inequality or zero-state observability of (2.118).

Analagous to Section 2.4.2, we now present a result due to van der Schaft [66] relating the solvability of the state feedback local nonlinear suboptimal  $H_\infty$  control problem for input-affine systems

$$\begin{aligned} \dot{x} &= a(x) + b(x)u + g(x)d, \quad a(0) = 0 \\ y &= x \\ z &= \begin{bmatrix} h(x) \\ u \end{bmatrix}, \quad h(0) = 0 \end{aligned} \tag{2.127}$$

to that of its linearization at zero

$$\begin{aligned} \dot{x} &= Ax + Bu + Gd \\ y &= x \\ z &= \begin{bmatrix} Hx \\ u \end{bmatrix} \end{aligned} \tag{2.128}$$

where  $A = \frac{\partial a}{\partial x}(0)$ ,  $B = b(0)$ ,  $G = g(0)$ , and  $H = \frac{\partial h}{\partial x}(0)$ .

**Theorem 2.4.1** *Consider (2.127) and (2.128) and assume  $(H, A)$  detectable. The following statements are equivalent.*

i.  $\exists$  a linear state feedback

$$u = Kx \tag{2.129}$$

such that the closed loop system (2.128), (2.129) is asymptotically stable and has induced  $L_2$  gain  $< \gamma$  (the linear state feedback  $H_\infty$  suboptimal control problem is solvable).

ii.  $\exists$  a solution  $X = X^T \geq 0$  to the Riccati equation

$$A^T X + XA + X\left(\frac{1}{\gamma^2}GG^T - BB^T\right)X + H^T H = 0 \tag{2.130}$$

such that all the eigenvalues of the matrix  $A - BB^T X + \frac{1}{\gamma^2}GG^T X$  have negative real parts.

iii.  $\exists$  a neighborhood  $W$  of 0, and a nonlinear state feedback

$$u = \alpha(x) \tag{2.131}$$

defined on  $W$  such that  $A + BK$ , with  $K = \frac{\partial \alpha}{\partial x}(0)$ , is asymptotically stable and the closed loop system (2.127), (2.131) has locally induced  $L_2$  gain  $< \gamma$  (the nonlinear local state feedback  $H_\infty$  suboptimal control problem is solvable on  $W$ ).

**Proof:** See [66]. ■

We note that if we replace the inequality with strict equality in (2.125), we obtain the so-called *Hamilton-Jacobi-Issacs* equation [67]. In solving the Hamilton-Jacobi-Issacs equation, we gain the added familiar interpretations for  $u_*$  and  $d_*$  as the minimizing control and worst-case disturbance, respectively. Also, in Theorem 2.4.1 above, technically we are speaking of solutions to the Hamilton-Jacobi-Issacs equation. A solution  $X$  to the Riccati equation in part ii implies the existence of a local solution  $V$  to the the Hamilton-Jacobi-Issacs equation for (2.127). Furthermore, it is true that

$$X = \frac{\partial^2 V}{\partial x^2}(0) \quad (2.132)$$

and that the linear part of (2.131) precisely equals (2.129) in part i. Thus, (2.129) solves the nonlinear state feedback suboptimal  $H_\infty$  control problem on some neighborhood of the origin. It is conjectured [67] that the domain of validity for (2.129) will always be smaller than for (2.131), but it has yet to be proven.

We conclude this section by making a connection between Hamilton-Jacobi-Issacs equations of nonlinear  $H_\infty$  control and Hamilton-Jacobi-Bellman equations for nonlinear optimal control. To do so, we observe that if we let  $\gamma \rightarrow \infty$ , the Hamilton-Jacobi-Issacs equation for (2.127) tends to the Hamilton-Jacobi-Bellman equation of optimal control for the system

$$\dot{x} = a(x) + b(x)u \quad (2.133)$$

with the cost functional

$$\int_0^\infty \left[ \frac{1}{2} h^T(x(t)) h(x(t)) + \frac{1}{2} u^T(t) u(t) \right] dt \quad (2.134)$$

i.e.,

$$V_x(x)a(x) - \frac{1}{2}V_x(x)b(x)b^T(x)V_x^T(x) + \frac{1}{2}h^T(x)h(x) = 0 \quad (2.135)$$

Local solutions for (2.135) always exist under the assumption of stabilizability of the linearized system [47]. This observation gives us a link between nonlinear  $H_\infty$  control and optimal nonlinear quadratic regulation which will be discussed further in Section 2.5.1.

We conclude this section with the statement that all the results presented in this section have dealt with the suboptimal  $H_\infty$  control problem. This is the case because complete theory for the optimal case does not yet exist. Recently, however, approximate techniques for computing the optimal induced  $L_2$  gain have appeared in the literature [35, 75].

## 2.5 State-Dependent Riccati Equation Techniques

This section presents the state-dependent Riccati equation (SDRE) theory of nonlinear regulation for input-affine systems as recently developed by Cloutier, D'Souza, and Mracek [13], and a nonlinear  $H_\infty$  solution technique based on similar ideas. The section on nonlinear regulation is taken with minor modification from [13], while the section on nonlinear  $H_\infty$  control presents a solution technique proposed in [13] combined with original work by this author. It is emphasized that much of the theory in Section 2.5.1 is recent and unverified. The main contribution of this dissertation is to verify the validity of the technique, which is accomplished in later sections.

### 2.5.1 The Nonlinear Regulator Problem

In the nonlinear regulator problem, we are interested in minimizing the infinite-horizon cost function

$$\text{minimize } J = \frac{1}{2} \int_{t_0}^{\infty} [x^T Q(x)x + u^T R(x)u] dt \quad (2.136)$$

subject to the nonlinear differential constraint

$$\dot{x} = a(x) + b(x)u \quad (2.137)$$

given state  $x \in \mathcal{R}^n$  and control  $u \in \mathcal{R}^m$ , with  $a, b, R, Q \in C^k$ ,  $k \geq 1$ , and where  $Q(x) = H^T(x)H(x) \geq 0$ , and  $R(x) > 0$  for all  $x$ . It is assumed that  $a(0) = 0$  so that the origin is an open

loop equilibrium point of the system. We seek stabilizing solutions of the form

$$u = L(x)x \quad (2.138)$$

where the nonlinear feedback gain  $L$  is a matrix function of  $x$ . The above formulation should be familiar from linear quadratic regulator (LQR) theory [41, 76] except that the matrices  $Q$ ,  $R$ , and  $L$  all have elements that are allowed to be functions of  $x$ . Thus, the cost function (2.136) may or may not be quadratic depending on whether  $Q$  and  $R$  are constant matrices or not.

The SDRE method hinges on being able to write the constraint dynamics (2.137) in a pointwise linear structure having state-dependent coefficient (SDC) form

$$\dot{x} = A(x)x + B(x)u \quad (2.139)$$

so that

$$a(x) = A(x)x \quad (2.140)$$

and

$$b(x) = B(x)$$

At this point it is important to note that the assumptions on the open loop drift function  $a$  guarantee that a global SDC parametrization of  $a$  exists [69]. One such factorization is [69]:

$$A(x) = \int_0^1 \nabla_x a(\lambda x) d\lambda \quad (2.141)$$

which is guaranteed to exist if  $a(0) = 0$  and  $a$  is  $C^1$ .

We now make the following definitions associated with the SDC form for later use.

- $\{H(x), A(x)\}$  is a **globally observable parametrization** of the nonlinear system if the pair  $\{H(x), A(x)\}$  is observable for all  $x$ .
- $\{A(x), B(x)\}$  is a **globally controllable parametrization** of the nonlinear system if the pair  $\{A(x), B(x)\}$  is controllable for all  $x$ .
- $\{H(x), A(x)\}$  is a **globally detectable parametrization** of the nonlinear system if the pair  $\{H(x), A(x)\}$  is detectable for all  $x$ .

- $\{A(x), B(x)\}$  is a **globally stabilizable parametrization** of the nonlinear system if the pair  $\{A(x), B(x)\}$  is stabilizable for all  $x$ .

We associate with the nonlinear cost function (2.136) the state-dependent (algebraic) Riccati equation (SDRE):

$$A^T(x)P(x) + P(x)A(x) - P(x)B(x)R^{-1}(x)B^T(x)P(x) + Q(x) = 0 \quad (2.142)$$

The SDRE nonlinear regulation technique is to solve (2.142), accepting only  $P(x) = P^T(x) \geq 0 \forall x$ , and construct the nonlinear feedback control by setting

$$u = -R^{-1}(x)B^T(x)P(x)x \quad (2.143)$$

so that the nonlinear feedback gain is

$$L(x) = -R^{-1}(x)B^T(x)P(x) \quad (2.144)$$

These equations can be solved analytically to produce an equation for each element of  $u$ , or solved numerically at a sufficiently high sampling rate, as was done in [53]. Note that quite clearly from (2.143), we must have full state feedback available in order to construct the control  $u$  using this method.

The local stability of the closed loop system resulting from using the SDRE nonlinear regulation technique is given by the following theorem from [13].

**Theorem 2.5.1** *In addition to  $a, b, R, Q \in C^k, k \geq 1$ , assume  $A \in C^k$  and  $A$  gives a globally stabilizable and detectable state-dependent coefficient parametrization of the nonlinear system. Then the SDRE nonlinear regulation control method has a closed loop solution which is locally asymptotically stable.*

**Proof:** The proof is repeated from [13], with only minor notational changes to be consistent with the remainder of this dissertation. Note that the closed loop solution is given by

$$\dot{x} = [A(x) - B(x)R^{-1}(x)B^T(x)P(x)]x \equiv F(x)x$$

and that the closed loop matrix function  $F(x)$  is guaranteed to be stable at every point  $x$  from standard Riccati equation theory. Under the  $C^k$  and stabilizable/detectable assumptions on the system parameters,  $P \in C^k$  so that  $F \in C^k$ . Expanding  $F$  in a partial Taylor series expansion about zero yields, for some neighborhood of the origin,

$$\dot{x} \approx F(0)x + \psi(x) \cdot \|x\|$$

where  $\psi(x)$  is of order  $k$  and

$$\lim_{\|x\| \rightarrow 0} \psi(x) = 0$$

In a smaller neighborhood about the origin, the linear term, which has a constant, stable coefficient matrix, dominates the higher-order term yielding local asymptotic stability. ■

At this point we remark that the proof of Theorem 2.5.1 relies on linearization arguments, and thus global stabilizability and detectability of the SDC parametrization are not required for local stability, but only stabilizability and detectability of the linearization of (2.137).

We now consider necessary conditions for optimality of the SDRE method. From the performance index and constrained dynamics we form the Hamiltonian function

$$\mathcal{H} = \frac{1}{2}x^T Q(x)x + \frac{1}{2}u^T R(x)u + \lambda^T [A(x)x + B(x)u] \quad (2.145)$$

with stationarity conditions

$$\mathcal{H}_u^T = 0 \quad (2.146)$$

$$\dot{\lambda} = -\mathcal{H}_x^T \quad (2.147)$$

$$\dot{x} = A(x)x + B(x)u \quad (2.148)$$

Using (2.143) and (2.145) we have

$$\mathcal{H}_u^T = R(x)u + B^T(x)\lambda \quad (2.149)$$

$$= R(x)[-R^{-1}(x)B^T(x)P(x)x] + B^T(x)\lambda \quad (2.150)$$

$$= B^T(x)[\lambda - P(x)x] \quad (2.151)$$

Thus,  $\mathcal{H}_u^T = 0$  if we choose

$$\lambda = P(x)x \quad (2.152)$$

Satisfying Equation (2.152) for all time will satisfy the  $\mathcal{H}_u$  optimality condition. From here we will drop the argument  $(x)$  notation for simplicity. Differentiating (2.152) with respect to time gives

$$\dot{\lambda} = \dot{P}x + P\dot{x} \quad (2.153)$$

Using the optimality condition (2.147) we also have

$$\dot{\lambda} = -Qx - \frac{1}{2}x^T Q_x x - \frac{1}{2}u^T R_x u - (x^T A_x^T + A^T + u^T B_x^T)\lambda \quad (2.154)$$

Equating (2.153) and (2.154) with substitutions from (2.137) and (2.143) gives

$$\dot{P}x + P(Ax - BR^{-1}B^T Px) = -Qx - \frac{1}{2}x^T Q_x x - \frac{1}{2}u^T R_x u - (x^T A_x^T + A^T + x^T PBR^{-1}B_x^T)Px \quad (2.155)$$

Rearrange to form

$$\dot{P}x + \frac{1}{2}x^T Q_x x + \frac{1}{2}u^T R_x u + x^T A_x^T Px - x^T PBR^{-1}B_x^T Px + [A^T P + PA - PBR^{-1}B^T P + Q]x = 0 \quad (2.156)$$

Furthermore, from (2.142) note that the term in brackets is the nonlinear regulation SDRE, which equals zero, and substituting for  $u$  one more time, (2.156) reduces to

$$\dot{P}x + \frac{1}{2}x^T Q_x x + \frac{1}{2}x^T PBR^{-1}R_x R^{-1}B^T Px + x^T A_x^T Px - x^T PBR^{-1}B_x^T Px = 0 \quad (2.157)$$

This is the *SDRE Necessary Condition for Optimality* which must be satisfied for the closed loop solution to be locally optimal. The authors of [13] note that the left hand side of (2.157) collapses to zero in the standard case of (infinite time) linear quadratic regulation, and thus this method is a true generalization of LQR theory.

The authors of [13] give the following theorem for the scalar case.

**Theorem 2.5.2** *Consider (2.136) and (2.137) for scalar  $x$ , i.e.,  $n = 1$ . Then there exists a unique SDC parametrization for  $x \neq 0$ :  $A(x) = a(x)/x$ , and the SDRE Necessary Condition for Optimality is always satisfied.*

**Proof:** The proof of uniqueness is trivial but repeated here from [13] for convenience. Let  $A_1(x)$  and  $A_2(x)$  be two SDC parametrizations of (2.137). Then  $a(x) = A_1(x)x = A_2(x)x$  so that  $[A_1(x) - A_2(x)]x = 0$ . Thus  $A_1(x) = A_2(x)$  for all  $x \neq 0$ . The proof of satisfaction of the SDRE Necessary Condition for Optimality is achieved by a straightforward but tedious algebraic exercise, and is not repeated here. ■

We remark that the uniqueness of  $A(x)$  in the scalar case could be extended to  $x = 0$  by assuming the continuity of  $A$ .

In the multistate case, it is stated in [13] that an infinite number of SDC parametrizations exists for (2.137). The proof relies on showing that at least two parametrizations exist, and that any linear combination of these two parametrizations is also a valid SDC parametrization. The existence of at least two parametrizations is trivial to show. The existence of an infinite number of parametrizations is more interesting to examine, and we therefore reproduce that part of the proof here. Thus, if we let  $A_1(x)$  and  $A_2(x)$  be two distinct SDC parametrizations, then  $a(x) = A_1(x)x = A_2(x)x$ . If we consider the hyperplane of SDC matrices  $A(x, \alpha) = \alpha A_1(x) + (1 - \alpha)A_2(x)$ , we find

$$\begin{aligned}
 A(x, \alpha)x &= \alpha A_1(x)x + (1 - \alpha)A_2(x)x \\
 &= \alpha a(x) + (1 - \alpha)a(x) \\
 &= \alpha a(x) + a(x) - \alpha a(x) \\
 &= a(x)
 \end{aligned} \tag{2.158}$$

and thus,  $A(x, \alpha)$  is a valid SDC parametrization for all  $\alpha \in \mathcal{R}$ , so that there exists an infinite number of parametrizations corresponding to the choice of  $\alpha$ . The authors of [13] then conjecture the existence of an optimal SDC parametrization leading to an optimal  $P(x, \alpha)$  solution of (2.142) and (2.157), and propose a procedure for finding it that results in a two-point boundary value problem to be solved. This author has observed that the hyperplane  $A(x, \alpha)$  above does not always span the space of valid SDC parametrizations, as it considers only two factorizations as a basis, and in the two-dimensional case, there may exist many more valid parametrizations which are not obtainable as linear combinations of a given pair of factorizations. Thus, this procedure and argument requires

closer examination. From [30] however, we are at least able to verify the conjectured existence of an optimal factorization, but only under certain assumed conditions.

We conclude this section by commenting that, although there are many theoretical issues to be resolved in the SDRE nonlinear regulator method, several impressive examples of solved design problems are presented in [14]. These examples show excellent agreement with results obtained using other well-established control techniques, thus suggesting utility in the SDRE nonlinear regulation method.

### 2.5.2 Nonlinear $H_\infty$ Control Via the State Feedback SDRE Method

In this section, we present proposed solution methods for the nonlinear  $H_\infty$  suboptimal control problem based on the SDRE technique. Solution approaches for input-affine systems under both state and output feedback are proposed in [13], without any proofs that the methods do indeed work. We present two original proofs for state feedback solution approaches to the nonlinear  $H_\infty$  suboptimal control problem based on the SDRE technique.

Consider the input-affine nonlinear system

$$\begin{aligned}\dot{x} &= a(x) + b(x)u + g(x)d \\ y &= x \\ z &= h(x) + d_{12}(x)u \\ d_{12}^T(x)[h(x) \ d_{12}(x)] &= [0 \ I]\end{aligned}\tag{2.159}$$

where  $x \in \mathcal{R}^n$ ,  $u \in \mathcal{R}^m$ ,  $d \in \mathcal{R}^r$ ,  $a_i(x) \in C^k(1 \leq i \leq n)$ ,  $b_{ij}(x) \in C^k(1 \leq i \leq n, 1 \leq j \leq m)$ ,  $g_{ij}(x) \in C^k(1 \leq i \leq n, 1 \leq j \leq r)$ ,  $k \geq 1$ , for which we desire to solve the local nonlinear  $H_\infty$  suboptimal control problem (see Section 2.4.1). Recall that  $z$  in (2.159) above is our penalized output, and the assumption in the last line of (2.159) requires separate penalties on states and controls and also control penalty scaling in order to simplify the problem. Note that  $d_{12}$  has replaced  $R$  as the control penalty matrix to reflect this scaling, and also to be consistent with the notation

used in the standard linear time-invariant developments [18]. In [13] the following SDRE-based solution approach to this problem is presented.

### SDRE Approach for Full State Feedback

- i. Parametrize (2.159) in SDC form

$$\begin{aligned}\dot{x} &= A(x)x + B(x)u + G(x)d \\ y &= x \\ z &= H(x)x + D_{12}(x)u\end{aligned}\tag{2.160}$$

- ii. Solve the  $H_\infty$  SDRE

$$A^T(x)P(x) + P(x)A(x) - P(x)[B(x)B^T(x) - \frac{1}{\gamma^2}G(x)G^T(x)]P(x) + H^T(x)H(x) = 0 \tag{2.161}$$

with  $\gamma$  sufficiently large so that the stability and complementarity properties (see Section 11.1) hold in order to obtain  $P(x) > 0 \forall x$ .

- iii. Construct the nonlinear  $H_\infty$  feedback control via

$$u(x) = -B^T(x)P(x)x \tag{2.162}$$

In order to prove the validity of an SDRE-based method, we need to strengthen the assumptions on (2.159) to agree with those of Section 2.4. This we now do as we formally state and prove our result.

**Theorem 2.5.3** *Consider (2.159) and assume  $z \in \mathcal{R}^s$ ,  $h(0) = 0$ . Also assume that all mappings in (2.159) are  $C^\infty$  and that  $\{H(0), A(0)\}$  is detectable and  $\{A(0), B(0)\}$  is stabilizable. Then the state feedback SDRE design procedure given by (2.161) and (2.162) yields a local solution to the nonlinear  $H_\infty$  control problem for (2.159).*

**Proof:** We consider the Hamiltonian point of view of solving the local nonlinear  $H_\infty$  suboptimal control problem for a fixed  $\gamma$ , as in Section 2.4.3. First, we note that (2.159) satisfies Assumption

2.4.1 in that all controls have an independent linear penalty placed on them. Thus, the theory of Section 2.4.3 may be applied to this problem, and our goal is to seek a  $C^1$ , (locally) positive semidefinite solution to the HJI (2.125). Recalling (2.119) we define the Hamiltonian

$$\mathcal{H}(x, V_x, d, u) = V_x(x)f(x, d, u) + \|z(x, u)\|^2 - \gamma^2 \|d\|^2 \quad (2.163)$$

Using (2.159), (2.163) becomes

$$\mathcal{H}(x, V_x, d, u) = V_x(x)[a(x) + b(x)u + g(x)d] + h^T(x)h(x) + u^T u - \gamma^2 d^T d \quad (2.164)$$

We propose a solution of the form  $V(x) = x^T P(x)x$ , with  $P(x) = P^T(x)$ . Using the proposed form for  $V$ , we have that

$$V_x(x) = 2x^T P(x) + x^T \left[ \frac{\partial P}{\partial x} x \right] \quad (2.165)$$

If we define the state-dependent matrix

$$M_P(x) \equiv \left[ \frac{\partial P}{\partial x} x \right] \quad (2.166)$$

and use the state-dependent parametrization (2.160), (2.164) can be written as:

$$\mathcal{H}(x, V_x, d, u) = x^T (2P + M_P)[Ax + Bu + Gd] + x^T H^T H x + u^T u - \gamma^2 d^T d \quad (2.167)$$

where we have dropped the  $(x)$  dependency notation for convenience. Now, the first-order necessary conditions (see (2.121))

$$\left( \frac{\partial \mathcal{H}}{\partial d} \right)_{d=d_*} = 0 = x^T (2P + M_P)G - 2\gamma^2 d_*^T \quad (2.168)$$

$$\left( \frac{\partial \mathcal{H}}{\partial u} \right)_{u=u_*} = 0 = x^T (2P + M_P)B + 2u_*^T \quad (2.169)$$

give

$$d_* = \frac{1}{2\gamma^2} G^T (2P + M_P)^T x \quad (2.170)$$

$$u_* = -\frac{1}{2} B^T (2P + M_P)^T x \quad (2.171)$$

Thus, the HJI we want to solve becomes

$$\begin{aligned} & x^T (2P + M_P) \left[ A - \frac{1}{2} B B^T (2P + M_P)^T + \frac{1}{2\gamma^2} G G^T (2P + M_P)^T \right] x + x^T H^T H x \\ & + \frac{1}{4} x^T (2P + M_P) B B^T (2P + M_P)^T x - \frac{1}{4\gamma^2} x^T (2P + M_P) G G^T (2P + M_P)^T x \leq 0 \end{aligned} \quad (2.172)$$

Using simple algebra and the fact that  $x^T Z x = x^T Z^T x$ , (2.172) simplifies to

$$\begin{aligned} x^T [A^T P + P A - P(BB^T - \frac{1}{\gamma^2} G G^T) P + H^T H] x + x^T [M_P A \\ - \frac{1}{4} M_P (B B^T - \frac{1}{\gamma^2} G G^T) M_P^T - M_P (B B^T - \frac{1}{\gamma^2} G G^T) P] x \leq 0 \end{aligned} \quad (2.173)$$

We observe that if we set the first bracketed term in (2.173) equal to zero, we obtain the  $H_\infty$  SDRE (2.161), and the remaining inequality

$$N(x) \equiv x^T [M_P A - \frac{1}{4} M_P (B B^T - \frac{1}{\gamma^2} G G^T) M_P^T - M_P (B B^T - \frac{1}{\gamma^2} G G^T) P] x \leq 0 \quad (2.174)$$

Now, since  $a$  and  $h$  were assumed to be  $C^\infty$ , they can be expanded, at least locally, in a Taylor series about the origin as

$$\begin{aligned} a(x) &= a(0) + \left( \frac{\partial a}{\partial x} \right)_{x=0} x + \frac{1}{2} x^T \left( \frac{\partial^2 a}{\partial x^2} \right)_{x=0} x + \cdots \\ h(x) &= h(0) + \left( \frac{\partial h}{\partial x} \right)_{x=0} x + \frac{1}{2} x^T \left( \frac{\partial^2 h}{\partial x^2} \right)_{x=0} x + \cdots \end{aligned} \quad (2.175)$$

and we recall that by assumption  $a(0) = 0$ ,  $h(0) = 0$ . Thus, recalling the SDC parametrization (2.160) we can write

$$\begin{aligned} A(x)x &= \left[ \left( \frac{\partial a}{\partial x} \right)_{x=0} + \frac{1}{2} x^T \left( \frac{\partial^2 a}{\partial x^2} \right)_{x=0} + \cdots \right] x \\ H(x)x &= \left[ \left( \frac{\partial h}{\partial x} \right)_{x=0} + \frac{1}{2} x^T \left( \frac{\partial^2 h}{\partial x^2} \right)_{x=0} + \cdots \right] x \end{aligned} \quad (2.176)$$

Thus,  $A(x)$  and  $H(x)$  are made up of polynomials, so that (2.161) implies that  $P$  will be made up of polynomial functions of  $x$ . The definition of polynomial equality requires all coefficients on like powers of  $x$  to be equal, so that from this discussion it is clear that, if  $P(x)$  solves (2.161), the constant coefficients must match and we must have

$$A^T(0)P(0) + P(0)A(0) - P(0)[B(0)B^T(0) - \frac{1}{\gamma^2}G(0)G^T(0)]P(0) + H^T(0)H(0) = 0 \quad (2.177)$$

Recalling the result of van der Schaft [66] at the end of Section 2.4.2, the solution  $P(x)$  of (2.161) has a 0th-order (constant) term  $P(0) \geq 0$  which, by assumption, is also a stabilizing solution of (2.177). Therefore, it must necessarily solve the local  $H_\infty$  suboptimal control problem for the *linearized system* about the origin. Furthermore, recall that any solution of the linearized problem is at least

a local solution of the nonlinear  $H_\infty$  suboptimal control problem. Thus, the control (2.171) locally solves the nonlinear  $H_\infty$  suboptimal control problem. A point to be made is that we have not addressed the inequality  $N \leq 0$ . This is because every term in this expression has at least one factor of the matrix function  $M_P$  in it. Recalling (2.166), we see that  $M_P(0) = 0$ , so that the 0th-order contribution of  $N$  to (2.173) is zero. Thus, by satisfying (2.161) we satisfy (2.177), and the HJI (2.173) is indeed satisfied to second order in  $x$  and thus is satisfied in some neighborhood of the origin, verifying that  $P(x)$  and  $u_*$  give a local solution to the nonlinear  $H_\infty$  suboptimal control problem, so that the theorem is proven. ■

Note that positive semidefinite  $P(0)$  is not as strong a requirement as in the state feedback approach of [13], which requires a positive definite solution  $P(x) \forall x$ , which is of course sufficient to guarantee  $P(0) \geq 0$ . We also note that, just as in the case of the SDRE nonlinear regulator, since the above arguments depend on linearized analysis, the global stabilizability and detectability assumptions on the SDC parametrizations have been relaxed to detectability of  $\{H(0), A(0)\}$  and stabilizability of  $\{A(0), B(0)\}$ .

Finally, recalling  $M_P(0) = 0$  and taking  $P(0)$  from (2.177), we see that the linear part of the optimal control (2.171) is  $u_* = -B^T(0)P(0)x$ . Thus, although (2.162) does not include the  $M_P$  term seen in (2.171), the linear parts of both controls are identical, and thus the full state feedback approach of Cloutier *et al.* [13] is valid if the proper strengthening of assumptions indicated above is performed.

An alternative approach to solving the nonlinear  $H_\infty$  control problem by the SDRE method may be derived not by proposing a form for the Lyapunov function  $V(x) = x^T P(x)x$ ,  $P(x) = P^T(x)$ , but instead proposing a form for the partial derivative (gradient) of the Lyapunov function,  $V_x(x) = 2x^T P(x)$ , with  $P(x)$  a symmetric matrix function of  $x$ . To see this, we first determine the HJI we wish to solve in terms of  $V_x(x)$ . Again dropping the  $(x)$  dependency for ease of notation, directly from (2.164) the first-order necessary conditions (see 2.121) give

$$\left( \frac{\partial \mathcal{H}}{\partial d} \right)_{d=d_*} = 0 = V_x G - 2\gamma^2 d_*^T \quad (2.178)$$

$$\left(\frac{\partial \mathcal{H}}{\partial u}\right)_{u=u_*} = 0 = V_x B + 2u_*^T \quad (2.179)$$

so that

$$d_* = \frac{1}{2\gamma^2} G^T V_x^T \quad (2.180)$$

$$u_* = -\frac{1}{2} B^T V_x^T \quad (2.181)$$

Substituting the above back into (2.164) and simplifying gives the HJI

$$V_x A x - \frac{1}{4} V_x B B^T V_x^T + \frac{1}{4\gamma^2} V_x G G^T V_x^T + x^T H^T H x \leq 0 \quad (2.182)$$

We now substitute the proposed form  $V_x = 2x^T P$  into (2.182) and simplify (again using  $x^T Z x = x^T Z^T x$ ) to obtain

$$x^T [A^T P + P A - P(B B^T - \frac{1}{\gamma^2} G G^T) P + H^T H] x \leq 0 \quad (2.183)$$

and we see comparing to (2.173), that we have eliminated the second bracketed term, leaving only the left hand side of the SDRE proposed in (2.161) that must be set less than or equal to zero to solve the nonlinear  $H_\infty$  problem. While it is significantly easier to solve the HJI (2.183) as compared to (2.173), the new wrinkle in this solution approach is that in the multistate case we must solve the set of simultaneous partial differential equations  $V_x(x) = 2x^T P(x)$  to obtain a Lyapunov function candidate  $V(x)$ , which we must then hope is at least locally positive semidefinite in order to guarantee local asymptotic stability. The conditions under which this may be done are explored in Sections 4.4 and 4.5. We thus have an apparent tradeoff between these two approaches. In the first method we have a Lyapunov function  $V(x) = x^T P(x) x$  which we can guarantee to be locally positive semidefinite if  $P(0)$  is a positive semidefinite matrix, and we have conditions (linearized stabilizability and detectability) that tell us when this will be the case. However, we have a difficult HJI to solve, involving two parts, the second of which contains coupled terms of the matrices  $P$  and  $M_P$ . In the second method, we need only solve the SDRE (2.161) to get a negative definite  $\dot{V}$ , but then it is much more difficult to deduce the existence of a positive semidefinite Lyapunov function  $V$ . An interesting observation is that, using this second approach, we always obtain the same optimal control as proposed by Cloutier *et al* [13]. A final comment that supports the second

proposed approach is the fact that as  $\gamma \rightarrow \infty$ , (2.182) and (2.183) (in the equality case) approach the correct Hamilton-Jacobi-Bellman equation (2.135) (properly accounting for the factor of  $1/2$  missing in (2.164)) and nonlinear regulator SDRE (2.161), as they should, while (2.173) does not. This second approach is thus consistent with linear  $H_\infty$  control approaches, which is appealing from an intuitive perspective. The relationship between these two design approaches is further explored in Section 4.5.

In this chapter we have reviewed relevant theory for stability analysis and control synthesis for nonlinear dynamic systems. Except for Section 2.5.2, in which we gave an original proof that the SDRE method does provide a local solution of the nonlinear  $H_\infty$  suboptimal control problem under suitable assumptions, all the material was compiled from existing references, to serve as background for the remainder of this dissertation.

### III. Motivational Example Problem

In this chapter, we apply the methods of Chapter 2 to an academic second-order single-input example problem. This application is intended to clarify the theory presented in Chapter 2, to highlight the differences between the various nonlinear control design methods, and to point to promising areas of research. The problem to be considered is taken from [34], and has several interesting aspects which will be pointed out as they are uncovered. In Section 3.1 we present the problem and perform some introductory analysis. In the remaining sections of the chapter we apply each of the four control methods of Chapter 2 in the same order they were presented there.

#### 3.1 Problem Setup and Introductory Analysis

Consider the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_1 x_2 \\ x_1^2 + d + u \end{bmatrix}, \quad z = \begin{bmatrix} x_2 - x_1^2 \\ u \end{bmatrix} \quad (3.1)$$

for which we immediately identify from our standard notation

$$\dot{x} = a(x) + b(x)u + g(x)d \quad (3.2)$$

that

$$a(x) = \begin{bmatrix} x_1 x_2 \\ x_1^2 \end{bmatrix}, \quad b(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = g(x) \quad (3.3)$$

Our control objective is to stabilize (3.1) while simultaneously attenuating the effect of the disturbance  $d$  on the penalized output  $z$ , assuming full state feedback. Thus, the problem we are trying to solve may be most naturally cast as a state feedback nonlinear  $H_\infty$  control problem, although the other methods of Chapter 2 may also be used at least to attempt stabilization. Stability and disturbance attenuation in the  $H_\infty$  sense was the premise of [34], in which the authors used this example problem for the stated reason that (3.1) cannot be stabilized by any linear full state feedback control law  $u = ax_1 + bx_2$ . This claim, if true, provides excellent motivation for employing nonlinear

control techniques, and also implies that the open loop system is itself unstable. We will now use the methods of Section 2.1 to verify this important claim.

In stability analysis of (3.1), we wish to first consider the unforced system (i.e., with  $d = 0$ ,  $u = 0$ ) without penalty variable:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_1 x_2 \\ x_1^2 \end{bmatrix} \quad (3.4)$$

Clearly, (3.4) has equilibrium points at  $x_1 = 0$  for all  $x_2 \in \mathcal{R}$ . Thus the  $x_2$ -axis of the  $x_1 x_2$  phase plane consists entirely of critical points. We shall now examine the stability of these critical points via the techniques of linearization and first integrals [68]. The Jacobian of (3.4) at any critical point  $(0, x_2)$  is easily computable as

$$J(0, x_2) = \left( \frac{\partial a}{\partial x} \right)_{(0, x_2)} = \begin{bmatrix} x_2 & x_1 \\ 2x_1 & 0 \end{bmatrix}_{(0, x_2)} = \begin{bmatrix} x_2 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.5)$$

Thus, each positive  $x_2$ -axis critical point has a 1-dimensional unstable manifold and a 1-dimensional center manifold, and each negative  $x_2$  axis critical point has a 1-dimensional stable manifold and a 1-dimensional center manifold. It is quite trivial to calculate the directions of the stable/unstable manifolds as parallel to the  $x_1$ -axis, and the direction of the center manifolds as the  $x_2$ -axis. The origin ( $x_1 = 0$ ,  $x_2 = 0$ ) clearly has a 2-dimensional center manifold. The existence of these center manifolds is one of the interesting aspects of this problem alluded to previously. As in [34], we are most interested in the behavior around the critical point at the origin. It turns out that (3.4) is separable and can be solved explicitly to obtain a **first integral** of the motion, which yields a parametrized set of trajectories which solutions of (3.4) must follow in the phase plane. This can be seen by dividing the  $\dot{x}_2$  equation by the  $\dot{x}_1$  equation to obtain

$$\frac{dx_2}{dx_1} = \frac{x_1^2}{x_1 x_2} = \frac{x_1}{x_2} \quad (3.6)$$

Rearranging gives

$$x_2 dx_2 = x_1 dx_1 \quad (3.7)$$

and integrating and multiplying by 2 yields

$$x_2^2 = x_1^2 + C, \quad C \in \mathcal{R} \quad (3.8)$$

Thus, solutions of (3.4) follow hyperbolic trajectories parametrized by the constant  $C$  which can be determined from the initial conditions. For zero initial conditions, (3.8) reduces to the degenerate hyperbola

$$x_2^2 = x_1^2$$

or equivalently to the orthogonal lines  $x_2 = \pm x_1$ , which are the center manifolds of the critical point at the origin. For any nonzero initial condition near the origin (arbitrarily close) which is not a critical point, the solution of (3.4) follows a hyperbolic trajectory to infinity, and is therefore unbounded, clearly indicating instability of the origin for the unforced system.

We now wish to consider how things change allowing only the presence of the control input  $u$ . We first consider the linearization of (3.1), and attempt to invoke Theorem 2.2.8. Through simple calculation one obtains the linear approximation  $\dot{x} = Ax + Bu$  to (3.1) as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad (3.9)$$

yielding the controllability matrix

$$M_c = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad (3.10)$$

which clearly has rank one. Thus, the linearization has one uncontrollable mode, corresponding to a zero eigenvalue in the  $A$  matrix. Thus, Theorem 2.2.8 provides no information as to whether or not (3.1) can be locally asymptotically stabilized by either linear or nonlinear feedback, and (3.1) represents a *critical system* for which we need a stronger analysis technique than the *Principle of Stability in the First Approximation*. We therefore attempt to verify the nonstabilizability of (3.1) via linear feedback by the Center Manifold Theorem, Theorem 2.1.7, and its associated Reduction Principle, Theorem 2.1.8.

Under an arbitrary linear full state feedback,  $u = ax_1 + bx_2$ ,  $a, b \in \mathcal{R}$ , (3.4) (with  $d = 0$ ) becomes

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_1 x_2 \\ x_1^2 + ax_1 + bx_2 \end{bmatrix} \quad (3.11)$$

which has Jacobian at the origin

$$J(0,0) = \begin{bmatrix} x_2 & x_1 \\ 2x_1 + a & b \end{bmatrix}_{(0,0)} = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix} \quad (3.12)$$

from which it can be seen that, for such a control to stabilize (3.1),  $b$  must be chosen negative. This may not be sufficient, however, as the behavior along the center manifold whose existence is implied by the zero eigenvalue of (3.12) may be unstable. Following the procedure outlined in Section 2.1.2, we can write (3.11) in the form of (2.21), (2.22) by identifying the matrix of right eigenvectors of (3.12) as

$$T = \begin{bmatrix} b & 0 \\ -a & 1 \end{bmatrix} \quad (3.13)$$

and performing the coordinate transformation

$$\begin{pmatrix} y \\ z \end{pmatrix} = T^{-1}x \quad (3.14)$$

to give

$$\begin{aligned} \dot{y} &= Ay + g(y, z) = by + ayz + z^2(b^2 - a^2) \\ \dot{z} &= Bz + h(y, z) = 0z + z(y - az) \end{aligned} \quad (3.15)$$

We now seek to solve the center manifold equation with  $y = \pi(z)$

$$\frac{\partial \pi(z)}{\partial z}(Bz + h(\pi(z), z)) = A\pi(z) + g(\pi(z), z) \quad (3.16)$$

or, making the appropriate associations from (3.15)

$$\frac{\partial \pi(z)}{\partial z}(z(\pi(z) - az)) = b\pi(z) + az\pi(z) + z^2(b^2 - a^2) \quad (3.17)$$

We solve (3.17) approximately by letting  $\pi(z) = \alpha z^2 + \mathcal{O}(z^3)$  and gathering like terms in powers of  $z$  to yield

$$\alpha = \frac{a^2 - b^2}{b} \quad (3.18)$$

Thus, the behavior along the center manifold is governed by the equation

$$\dot{z} = z \left( \frac{a^2 - b^2}{b} z^2 + \mathcal{O}(z^3) - az \right) = -az^2 + \mathcal{O}(z^3) \quad (3.19)$$

Invoking Lemmas 2.1.2 and 2.1.3, we conclude that (3.19) is unstable for all values of  $a$ , and thus by Theorem 2.1.8 indeed we find that (3.1) is not asymptotically stabilizable by any linear full state feedback  $u = ax_1 + bx_2$ . Thus, we have well motivated the need to seek nonlinear feedbacks to stabilize (3.1), which we now proceed to do.

### 3.2 Feedback Linearization

In this section we attempt to apply the techniques of Section 2.2 to (3.1). In particular, since (3.1) is a single-input system, we will be drawing heavily on the material of Section 2.2.2. We first consider the State Space Exact Linearization Problem for (3.1), since if we can solve it, we can at least stabilize the origin. Recall that in the terminology of Section 2.2.2 we have  $b(x) = [0 \ 1]^T$  and  $n = 2$ , and from Theorem 2.2.3 that the problem is solvable in a neighborhood of the origin if the following two conditions hold:

- i. the matrix  $[b(0) \ ad_a b(0)]$  has rank 2
- ii. the distribution  $D = \text{span}\{b\}$  is involutive near  $x = 0$ .

Using the definition of the Lie bracket (2.36), we find

$$[a, b] = \frac{\partial b}{\partial x} a(x) - \frac{\partial a}{\partial x} b(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 x_2 \\ x_1^2 \end{bmatrix} - \begin{bmatrix} x_2 & x_1 \\ 2x_1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -x_1 \\ 0 \end{bmatrix} \quad (3.20)$$

and

$$[b, ad_a b] = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (3.21)$$

From (3.20) we see that the matrix in item i above is

$$\begin{bmatrix} 0 & -x_1 \\ 1 & 0 \end{bmatrix}_{(0,0)} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad (3.22)$$

which clearly does not have rank 2, and thus the SISO State Space Exact Linearization Problem is technically not solvable for (3.1) at the origin. Recalling (2.43), we see that the distribution  $D$  is indeed involutive, however, so that item ii is satisfied. The failure of item i above is associated with the lack of existence of a suitable output function for which (3.1) has a well-defined relative degree,  $r = 2$ , at the point  $(0, 0)$ , which in turn implies a lack of controllability when  $x_1 = 0$ . From (3.22) we see that, if we consider the problem at any point for which  $x_1 \neq 0$ , we satisfy item i. Since every point with  $x_1 = 0$  is an equilibrium, however, it may not be crucial to control the system at those points. Thus, although item i above is not satisfied, we will proceed with the design procedure in hopes of obtaining a stabilizing control. We seek a scalar output function  $y = c(x)$  such that

$$L_b c = \frac{\partial c}{\partial x} b = 0 \quad (3.23)$$

$$L_{ad_a} b c \neq 0. \quad (3.24)$$

Equation (3.23) implies

$$\begin{bmatrix} \frac{\partial c}{\partial x_1} & \frac{\partial c}{\partial x_2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\partial c}{\partial x_2} = 0 \Rightarrow c = c(x_1) \quad (3.25)$$

so that  $c$  is a function of  $x_1$  only. Equation (3.24) implies

$$\begin{bmatrix} \frac{\partial c}{\partial x_1} & \frac{\partial c}{\partial x_2} \end{bmatrix} \begin{bmatrix} -x_1 \\ 0 \end{bmatrix}_{(0,0)} = -x_1 \frac{\partial c}{\partial x_1} \Big|_0 \neq 0 \quad (3.26)$$

which can be solved by (arbitrarily) letting

$$\frac{\partial c}{\partial x_1} = \frac{1}{x_1}$$

so that

$$c(x_1) = \ln(x_1). \quad (3.27)$$

Taking our new coordinates as in Lemma 2.2.3, we get

$$z_1 = c(x_1) = \ln(x_1) \quad (3.28)$$

$$z_2 = L_a c(x_1) = \begin{bmatrix} \frac{1}{x_1} & 0 \end{bmatrix} \begin{bmatrix} x_1 x_2 \\ x_1^2 \end{bmatrix} = x_2 \quad (3.29)$$

and we see that this transformation is well-defined only for  $x_1 > 0$ . Although we do not have a coordinate transformation valid in a neighborhood of the origin, we will proceed with the design steps to illustrate the theory, and consider possible modification of the control scheme to account for this problem later. Recall that the linearizing feedback can be found via

$$u = \phi(x) + \theta(x)v \quad (3.30)$$

with

$$\phi(x) = \frac{-L_a^2 c(x)}{L_b L_a c(x)} \quad (3.31)$$

$$\theta(x) = \frac{1}{L_b L_a c(x)} \quad (3.32)$$

Computing the necessary values we find

$$L_a^2 c = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 x_2 \\ x_1^2 \end{bmatrix} = x_1^2 \quad (3.33)$$

$$L_b L_a c = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \quad (3.34)$$

so that the linearizing feedback is

$$u = -x_1^2 + v \quad (3.35)$$

and we see that this feedback control is well-defined for all  $x$  despite the above-mentioned restricted region of applicability of the coordinate transformation. In the transformed set of coordinates, (3.1)

becomes

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v \quad (3.36)$$

The transformed system (3.36) has clearly been rendered linear by the feedback (3.35), and a simple calculation gives a controllability matrix of

$$M_c = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (3.37)$$

so that (3.36) is clearly seen to be controllable. Now, if the coordinate transformation were valid in a neighborhood of the origin, we could, for example, choose  $v = az_1 + bz_2$  with  $b$  negative and large enough and  $a$  negative and small enough so that the eigenvalues of the closed loop system

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ a & b \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \quad (3.38)$$

given by

$$\lambda = \frac{1}{2} \left[ b \pm (b^2 + 4a)^{\frac{1}{2}} \right] \quad (3.39)$$

are both negative, thus ensuring stability of the closed loop system. Since the coordinate transformation is not locally valid, however, we cannot guarantee closed loop stability using this control. When the coordinate transformation (3.27) is undefined (when  $x_1 \leq 0$ ), it might be possible to modify the above control scheme to provide stability, but this would probably require some ad hoc techniques such as switching logic based on the sign of  $x_1$  in the controller, and criteria to tell the controller not to try to evaluate  $z_1$  when  $x_1 = 0$ . These considerations are beyond the scope of feedback linearization theory, and will therefore not be pursued. Alternatively, recall from (3.22) that the conditions for solvability of the SISO State Space Exact Feedback Linearization Problem are satisfied at any point for which  $x_1 \neq 0$ . Following the above steps, it is easy to show that if we consider a point  $(x_1, x_2) = (k, 0)$ ,  $k \neq 0$ , by defining our output function as  $c(x) = x_1$ , (3.1) can be transformed via the valid (in a neighborhood of  $(x_1, x_2) = (k, 0)$ ) coordinate transformation  $(z_1, z_2) = (x_1, x_1 x_2)$  to (3.36) above. Thus, we could create and locally stabilize a critical point somewhere close to the origin, but our neighborhood of attraction would never include the  $x_2$ -axis. In summary, from the above discussion we have seen that not meeting the technical requirements for exact state space feedback linearization of (3.1) effectively prevents us from using feedback linearization theory to provide a control scheme for stabilizing the origin of (3.1) as desired.

Accepting this, for tutorial purposes we nevertheless suppose that we have obtained a valid coordinate transformation, and proceed with the disturbance attenuation part of the problem. In the transformed state space we find that our system becomes

$$\begin{bmatrix} \dot{w}_1 \\ \dot{w}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} d + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v \quad (3.40)$$

$$z = \begin{bmatrix} w_2 - e^{2w_1} \\ -e^{2w_1} + v \end{bmatrix} \quad (3.41)$$

and we see that, even though we have assumed a solution to the SISO State Space Exact Linearization Problem, the input-output behavior of our system remains highly nonlinear. Thus, we cannot employ linear  $H_\infty$  control theory directly to solve the problem. We could, however, employ linear  $H_\infty$  control theory on the linearization of (3.41) (with (3.40) unchanged since it is already linear) to obtain a local solution to the problem, as per Theorem 2.4.1. To do so we must solve the state feedback algebraic Riccati equation

$$A^T X + X A + X \left( \frac{1}{\gamma^2} G G^T - B B^T \right) X + H^T H = 0 \quad (3.42)$$

where from the linearization of (3.41) we have

$$H = \begin{bmatrix} -2 & 1 \\ -2 & 0 \end{bmatrix}$$

Letting

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix}$$

and substituting the appropriate values, (3.42) becomes

$$\begin{aligned} & \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} + \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ & + \lambda \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} + \begin{bmatrix} -2 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.43) \end{aligned}$$

where, as in [34], we have defined

$$\lambda = \left( \frac{1}{\gamma^2} - 1 \right)$$

For illustrative purposes, we shall solve (3.42) for two values of  $\lambda$ . It can be seen from (3.43) that we need  $\gamma \geq 1$  to be guaranteed existence of a stabilizing solution. We shall therefore consider the suboptimal  $H_\infty$  control problem for  $\gamma^2 = \frac{100}{99}$  (*alternatively*, 2) so that we have correspondingly  $\lambda = -\frac{1}{100}$  (*alternatively*,  $-\frac{1}{2}$ ). Note that  $\gamma^2 = \frac{100}{99}$  is closer to the optimal value of one than is  $\gamma^2 = 2$ . Simplifying (3.43) results in three nonlinear equations to solve simultaneously:

$$\begin{aligned} \lambda x_{12}^2 + 8 &= 0 \\ 2x_{12} + \lambda x_{22}^2 + 1 &= 0 \\ x_{11} + \lambda x_{12}x_{22} - 2 &= 0 \end{aligned} \tag{3.44}$$

For  $\lambda = -\frac{1}{2}$ , (3.44) can be solved to give the stabilizing solution of (3.43)

$$X = \begin{bmatrix} 2 + 6\sqrt{2} & 4 \\ 4 & 3\sqrt{2} \end{bmatrix} \tag{3.45}$$

which yields the control

$$v = -B^T X w = -4w_1 - 3\sqrt{2}w_2 \tag{3.46}$$

Thus, the control for the original system for  $\lambda = -\frac{1}{2}$  is given by

$$u = -x_1^2 - 4\ln(x_1) - 3\sqrt{2}x_2 \tag{3.47}$$

For  $\lambda = -\frac{1}{100}$ , we obtain the stabilizing solution of (3.43)

$$X = \begin{bmatrix} 3.2 & 20\sqrt{2} \\ 20\sqrt{2} & 75.87 \end{bmatrix} \tag{3.48}$$

which yields the control

$$v = -B^T X w = -20\sqrt{2}w_1 - 75.9w_2 \tag{3.49}$$

so that  $u$  for the original system for  $\lambda = -\frac{1}{100}$  is given by

$$u = -x_1^2 - 20\sqrt{2}\ln(x_1) - 75.9x_2 \tag{3.50}$$

Examination of (3.47) and (3.50) shows that neither control is defined for  $x_1 \leq 0$ , as expected from our previous discussion, and also that the near-optimal solution (for  $\lambda = -\frac{1}{100}$ ) requires significantly greater control effort as we might expect, since we are attempting a greater level of disturbance attenuation.

Thus, assuming we had a valid coordinate transformation, state space exact feedback linearization would allow us to solve the  $H_\infty$  suboptimal control problem at least locally. This result is significant because, if we had tried to solve the  $H_\infty$  suboptimal control problem for the original system via its linearization, we would obtain  $u = -\sqrt{2}x_2$ , and thus would have failed to stabilize the system since we have already shown that (3.1) cannot be stabilized by any linear feedback. We will say more about the solution given above later, in comparing it to solutions obtained by other design methods.

Although the solution above would be a valid local solution to the nonlinear  $H_\infty$  suboptimal control problem, ideally we would like to solve this problem by completely feedback linearizing the nonlinear input-output behavior, and then using linear  $H_\infty$  theory on the resulting system to obtain a larger region of validity for the control than what we get by using the linearization of the output equation. Thus, we now seek a solution to the input-output feedback linearization problem for (3.1).

Recall from Section 2.2.2 that a sufficient condition for a solution to exist to our input-output linearization problem for a SISO system with specified output is that the system have relative degree  $r = n$  for the specified output. This is equivalent to the specified output solving the state space exact linearization problem, which  $z = x_2 - x_1^2$  clearly does not, since it does not solve (3.26). We know this is sufficient because of Theorem 2.4.1. From Section 2.2 it is clear that, if for the given  $z$  we have some well-defined relative degree  $r$ , we can obtain a linearized input-output relationship. Further, if the zero dynamics of the transformed system are LAS, we could use any stabilizing control on the linearized subsystem (including one from linear  $H_\infty$  theory) to guarantee stability of the overall closed loop system. It is also true that if we did use the linear  $H_\infty$  control, the closed loop system would have the desired induced  $L_2$  gain properties, since the zero dynamics of the transformed system are totally unobservable in the linearized subsystem, and the output of interest depends only

on the linearized subsystem. This point turns out to be moot for the example problem, however, since the output  $z = x_2 - x_1^2$  yields a relative degree one system with unstable zero dynamics, as shown below.

A simple calculation, letting  $y_1 = h = x_2 - x_1^2$  yields

$$L_b h = [-2x_1 \ 1] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \quad (3.51)$$

thus indicating the system has relative degree one. Also, we find

$$L_a h = [-2x_1 \ 1] \begin{bmatrix} x_1 x_2 \\ x_1^2 \end{bmatrix} = x_1^2 - 2x_1^2 x_2 \quad (3.52)$$

so that the linearizing control is

$$u = \frac{1}{L_a h} (-L_a h + v) = 2x_1^2 x_2 - x_1^2 + v \quad (3.53)$$

In the transformed coordinates we obtain the input-output linearized system

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \begin{bmatrix} v \\ y_2(y_1 + y_2^2) \end{bmatrix} \quad (3.54)$$

The zero dynamics of (3.54) are obtained by requiring  $y_1 = 0$ ,  $\dot{y}_1 = 0$  so that  $v = 0$ , and leaving

$$\dot{y}_2 = y_2^3 \quad (3.55)$$

From Lemma 2.1.3 it is trivial to conclude instability of (3.55).

### 3.3 Recursive Backstepping

In this section, we apply the theory of Section 2.3 to (3.1). As in the previous section, we first start with stabilizing designs. We then seek ways to tailor the algorithm to achieve the desired disturbance attenuation objective.

Recall (3.1), and observe that without considering the output  $z$  this system is already in the desired configuration for a backstepping design. Thus, we consider the first scalar equation in (3.1),

$$\dot{x}_1 = x_1 x_2 \quad (3.56)$$

for which the state variable  $x_2$  makes a logical choice for a pseudocontrol. Under recursive backstepping, our objective is to choose a suitable desired value for  $x_2$ , which we call a stabilizing function,  $\alpha(x_1)$ , such that if we can achieve that value perfectly, then (3.56) will be stabilized. One logical choice for  $\alpha(x_1)$  above would be

$$\alpha(x_1) = -x_1^2 \quad (3.57)$$

so that the closed loop system equation would be

$$\dot{x}_1 = -x_1^3 \quad (3.58)$$

which can trivially be shown to be GAS by using the Lyapunov function  $V(x_1) = \frac{1}{2}x_1^2$  and the arguments of Section 2.1.1, or simply by invoking Lemma 2.1.3. However, we recognize that we will never achieve (3.57) perfectly by choosing  $u$  in the second scalar equation of (3.1),

$$\dot{x}_2 = x_1^2 + u \quad (3.59)$$

so we introduce the error variable  $e = x_2 - \alpha(x_1) = x_2 + x_1^2$ , and in transformed coordinates obtain

$$\begin{aligned} \dot{x}_1 &= x_1(e - x_1^2) \\ \dot{e} &= x_1^2 - 2x_1^4 + 2x_1^2e + u + d \end{aligned} \quad (3.60)$$

We now consider the globally positive definite Lyapunov function  $V(x_1, e) = \frac{1}{2}x_1^2 + \frac{1}{2}e^2$ , differentiating with respect to time to obtain (letting  $d = 0$ )

$$\begin{aligned} \dot{V} &= x_1\dot{x}_1 + e\dot{e} \\ &= x_1^2(e - x_1^2) + e(x_1^2 - 2x_1^4 + 2x_1^2e + u) \\ &= -x_1^4 + e(2x_1^2 + 2x_1^2e - 2x_1^4 + u) \end{aligned} \quad (3.61)$$

Thus, choosing

$$u = 2x_1^4 - 2x_1^2e - 2x_1^2 - c_1e = -2x_1^2 - 2x_1^2x_2 - c_1(x_2 + x_1^2) \quad (3.62)$$

with  $c_1 > 0$  guarantees the closed loop system is GAS, and that  $x_1$  and  $e$  converge asymptotically to zero as time goes to infinity.

Alternatively, one might choose the stabilizing function above as

$$\alpha(x_1) = -1 \quad (3.63)$$

so that in the closed loop (3.56) would become

$$\dot{x}_1 = -x_1 \quad (3.64)$$

which is not only stable, but stable in the first approximation. Following the same procedure as above, for this choice of  $\alpha(x_1)$  we obtain

$$\begin{aligned} \dot{V} &= x_1 \dot{x}_1 + e \dot{e} \\ &= x_1^2(e - 1) + e(x_1^2 + u) \\ &= -x_1^2 + e(2x_1^2 + u) \end{aligned} \quad (3.65)$$

Thus, now we could choose the control

$$u = -2x_1^2 - c_1 e = -2x_1^2 - c_1(x_2 + 1) \quad (3.66)$$

with  $c_1 > 0$  to guarantee the closed loop system is GAS, and the same asymptotic convergence properties for  $x_1$  and  $e$  hold. However, we must be careful here, because since  $e = x_2 + 1$ , by making  $(x_1, e) = (0, 0)$  GAS, we have rendered the equilibrium  $(x_1, x_2) = (0, -1)$  GAS in the original system coordinates, which was not our objective. This could have posed problems if we had only been trying to obtain local asymptotic stability, so that in view of (3.65) we had chosen  $u = -c_1 e = -c_1(x_2 + 1)$ . We know from Section 3.1 that such a linear control fails to stabilize  $(x_1, x_2) = (0, 0)$ , although it does make  $(x_1, e) = (0, 0)$  LAS. Nevertheless, we see from (3.61) that we could have chosen  $u$  according to (3.66) as a result of our first choice of stabilizing function, simply choosing  $c_1$  large enough for the lower-order negative definite terms to dominate the higher-order cross terms in a neighborhood of the origin. Comparing (3.62) and (3.66), we see that (3.66) is a much simpler and, in general, smaller magnitude control, making it appear a more attractive choice.

It is clear from this discussion that recursive backstepping can successfully be used to globally stabilize (3.1). The above application of the theory also clearly points out the nonuniqueness of

solutions mentioned in Section 2.3, and thus the ensuing need to exercise sound engineering judgment in using this method. As shown above, the obvious choice for the control is not always the best choice, and one principle to keep in mind is to avoid unnecessary cancellations whenever possible.

Recursive backstepping is aimed primarily at stabilization, and thus we note that we have yet to address the disturbance attenuation part of the design problem, which we now attempt to do. The algorithm itself provides no direct means of achieving this sort of objective, and we therefore present an original approach to solving the problem. Recall from the above discussion that  $x_1(t)$  and  $e(t)$  are regulated to zero as time goes to infinity. Thus, we might achieve the desired disturbance attenuation objective by choosing  $e(t)$  to be equal to the quantity we desire to make small, i.e., set  $e = z = x_2 - x_1^2$ . We then get

$$\begin{aligned}\dot{x}_1 &= x_1(e + x_1^2) \\ \dot{e} &= x_1^2 - 2x_1^4 - 2x_1^2e + u + d\end{aligned}\tag{3.67}$$

for the system equations in  $x, e$  coordinates. Note that in so doing we are implicitly choosing a stabilizing function  $\alpha(x_1) = x_1^2$ , which actually makes (3.56) unstable in the closed loop! This is not a standard choice of stabilizing function, and since it makes the first equation of (3.67) unstable, it prevents us from sequentially building up a simple, suitable Lyapunov function as is done in the normal procedure. For if we choose  $V$  as before we find

$$\begin{aligned}\dot{V} &= x_1\dot{x}_1 + e\dot{e} \\ &= x_1^2(e + x_1^2) + e(x_1^2 - 2x_1^4e - 2x_1^2e + u) \\ &= x_1^4 + e(2x_1^2 - 2x_1^2e - 2x_1^4 + u)\end{aligned}\tag{3.68}$$

and we see that no choice of  $u$  can render the  $x_1^4$  term negative definite. However, we see from (3.67) that we can obtain an  $x_1^4$  term in  $\dot{V}$  by including an  $x_1^2e$  cross term in  $V$ . Thus, we propose the candidate Lyapunov function

$$V(x_1, e) = \frac{1}{2}ax_1^2 + \frac{1}{2}be^2 + cx_1^2e, \quad a, b > 0\tag{3.69}$$

which will be locally positive definite in a neighborhood of the origin for any  $c \in \mathcal{R}$ . Differentiating and grouping like terms yields

$$\dot{V} = -2cx_1^6 - 2bx_1^4e + (a+c)x_1^4 + 2(c-b)x_1^2e^2 + (a+b)x_1^2e + u(be + cx_1^2) \quad (3.70)$$

from which we see we need  $c > 0$  for negative definiteness of the  $x_1^6$  term. We also see a third-order cross term  $x_1^2e$  that we cannot dominate since we have no  $x_1^2$  term, leading us to consider a control of the form

$$u = -c_1e + dx_1^2, \quad c_1 > 0 \quad (3.71)$$

Using this control, (3.70) becomes

$$\dot{V} = -2cx_1^6 - 2bx_1^4e + (a+c+cd)x_1^4 + 2(c-b)x_1^2e^2 + (a+b+bd-cc_1)x_1^2e - bc_1e^2 \quad (3.72)$$

and we now can choose coefficients to eliminate the undesirable  $x_1^2e$  cross term. To ensure local negative definiteness of (3.72), we want to choose the constants such that

$$a + b + bd - cc_1 = 0 \quad (3.73)$$

$$a + c + cd \leq 0 \quad (3.74)$$

and  $a$ ,  $b$ ,  $c$ , and  $c_1$  are all positive. These conditions are sufficient to guarantee local negative definiteness of (3.72) because, for any choice of coefficients such that  $c - b$  is finite and positive, there exists some neighborhood of the origin such that the  $x_1^2e^2$  cross term in (3.72) will be dominated by the  $e^2$  and  $x_1^4$  terms, provided they have negative coefficients. One set of choices that meets the above conditions is

$$a = 1, \quad b = \frac{1}{36}, \quad c = \frac{1}{8}, \quad d = -10, \quad c_1 = 6 \quad (3.75)$$

Using these coefficients, our final Lyapunov function becomes

$$V(x_1, e) = \frac{1}{2}x_1^2 + \frac{1}{72}e^2 + \frac{1}{8}x_1^2e \quad (3.76)$$

which has derivative

$$\dot{V} = -\frac{1}{4}x_1^6 - \frac{1}{18}x_1^4e - \frac{1}{8}x_1^4 + \frac{7}{36}x_1^2e^2 - \frac{1}{6}e^2 \quad (3.77)$$

The control is

$$u = -6e - 10x_1^2 = -4x_1^2 - 6x_2 \quad (3.78)$$

which, as discussed earlier, is guaranteed to locally stabilize (3.1) and also cause local asymptotic convergence of  $x_1$  and  $e = x_2 - x_1^2$  to zero as time goes to infinity.

### 3.4 Nonlinear $H_\infty$ Control

In this section, we apply the state feedback nonlinear  $H_\infty$  control theory of Section 2.4.3 to (3.1). We note first of all that (3.1) satisfies Assumption 2.4.1, since the rank of  $\partial z / \partial u = [0 \ 1]^T$  equals one, which is the number of inputs for this SISO problem. Thus, we proceed to apply the method by defining the Hamiltonian

$$\mathcal{H} = V_x f(x, u, d) + z^T z - \gamma^2 d^T d = V_{x_1} x_1 x_2 + V_{x_2} (x_1^2 + u + d) + (x_2 - x_1^2)^2 + u^2 - \gamma^2 d^2 \quad (3.79)$$

The first-order necessary conditions (see 2.121)

$$\left( \frac{\partial \mathcal{H}}{\partial d} \right)_{d=d_*} = 0 = V_{x_2} - 2\gamma^2 d_* \quad (3.80)$$

$$\left( \frac{\partial \mathcal{H}}{\partial u} \right)_{u=u_*} = 0 = V_{x_2} + 2u_* \quad (3.81)$$

give the worst case disturbance and optimal control, respectively, as

$$d_* = \frac{1}{2\gamma^2} V_{x_2} \quad (3.82)$$

$$u_* = -\frac{1}{2} V_{x_2} \quad (3.83)$$

Substituting (3.82) and (3.83) into (3.79), we wish to find a positive semidefinite solution to the Hamilton-Jacobi Inequality (see (2.125))

$$\mathcal{H}_* = V_{x_1} x_1 x_2 + V_{x_2} (x_1^2 - \frac{1}{2} V_{x_2} + \frac{1}{2\gamma^2} V_{x_2}) + x_2^2 - 2x_1^2 x_2 + x_1^4 + \frac{1}{4} V_{x_2}^2 - \frac{1}{4\gamma^2} V_{x_2}^2 \leq 0 \quad (3.84)$$

which simplifies to

$$V_{x_1} x_1 x_2 + V_{x_2} x_1^2 + x_2^2 - 2x_1^2 x_2 + x_1^4 + \frac{1}{4} V_{x_2}^2 \left( \frac{1}{\gamma^2} - 1 \right) \leq 0 \quad (3.85)$$

If, as in Section 3.3, we let  $\lambda = \frac{1}{\gamma^2} - 1$ , then (3.85) can be written

$$V_{x_1}x_1x_2 + V_{x_2}x_1^2 + x_2^2 - 2x_1^2x_2 + x_1^4 + \frac{\lambda}{4}V_{x_2}^2 \leq 0 \quad (3.86)$$

If we propose a solution of the form  $V = ax_1^2 + bx_2^2$ ,  $a, b > 0$ , we find (3.86) becomes

$$x_1^4 + (2a + 2b - 2)x_1^2x_2 + (1 + \lambda b^2)x_2^2 \leq 0 \quad (3.87)$$

and we see that we can not solve (3.87) in a neighborhood of the origin due to the positive coefficient on  $x_1^4$ . We therefore propose a candidate solution of the form  $V = ax_1^2 + bx_2^2 + cx_1^2x_2$ ,  $a, b > 0$ , for which (3.86) becomes

$$(1 + c + \frac{\lambda c^2}{4})x_1^4 + 2cx_1^2x_2^2 + (2a + 2b - 2 + \lambda bc)x_1^2x_2 + (1 + \lambda b^2)x_2^2 \leq 0 \quad (3.88)$$

From the  $x_2^2$  coefficient we clearly must have  $\lambda < 0$  in order to ensure local solution of (3.88), which implies  $\gamma > 1$ . Thus, as in Section 3.3, we will consider two cases: a near-optimal case where  $\lambda = -\frac{1}{100}$  corresponding to  $\gamma^2 = \frac{100}{99}$ , and a more suboptimal case where  $\lambda = -\frac{1}{2}$  corresponding to  $\gamma^2 = 2$ . Local solutions to (3.88) are not unique, consisting of any set of coefficients such that

$$2a + 2b - 2 + \lambda bc = 0 \quad (3.89)$$

$$1 + c + \frac{\lambda c^2}{4} \leq 0 \quad (3.90)$$

$$1 + \lambda b^2 \leq 0 \quad (3.91)$$

and  $a$  and  $b$  are positive. For  $\lambda = -\frac{1}{100}$ , one such set of coefficients is

$$a = 16, b = 10, c = 500 \quad (3.92)$$

and we note that the large positive  $c$  term is required to satisfy (3.89) and (3.90) ((3.89) requires  $c$  to be positive, and  $c = 400$  is the smallest positive  $c$  satisfying (3.90)). The corresponding expressions for  $V$ ,  $\mathcal{H}_*$ , and the control  $u$  are

$$\begin{aligned} V &= 16x_1^2 + 10x_2^2 + 500x_1^2x_2 \\ \mathcal{H}_* &= -124x_1^4 + 1000x_1^2x_2^2 - .44x_2^2 \\ u &= -10x_2 - 250x_1^2 \end{aligned} \quad (3.93)$$

For  $\lambda = -\frac{1}{2}$ , one such set of coefficients is

$$a = 16, b = 10, c = 10 \quad (3.94)$$

and we note that we do not need nearly as large a  $c$  term to satisfy (3.89) and (3.90). The corresponding expressions for  $V$ ,  $\mathcal{H}_*$ , and the control  $u$  are

$$\begin{aligned} V &= 16x_1^2 + 10x_2^2 + 10x_1^2x_2 \\ \mathcal{H}_* &= -\frac{3}{2}x_1^4 + 20x_1^2x_2^2 - 49x_2^2 \\ u &= -10x_2 - 5x_1^2 \end{aligned} \quad (3.95)$$

Comparing (3.93) with (3.95), we see that the near-optimal solution requires much greater control effort in terms of gain on  $x_1^2$ , as we might expect since we are trying to achieve a higher level of disturbance attenuation. However, the increase in disturbance attenuation (reduction in induced  $L_2$  gain) is only about a factor of two, while the required gain increase is near a factor of *fifty*. We thus seem to be seeing the familiar type of high bandwidth phenomenon we see in linear  $H_\infty$  control as we approach the optimal solution.

For later comparison purposes, it is also of interest to compute the solutions to the Hamilton-Jacobi-Issacs equations for the above cases. This involves solving (3.90) and (3.91), in addition to (3.89), with equality. For  $\lambda = -\frac{1}{100}$ , the solution is

$$a = 11, b = 10, c = 400 \quad (3.96)$$

giving

$$\begin{aligned} V &= 11x_1^2 + 10x_2^2 + 400x_1^2x_2 \\ \mathcal{H}_* &= \mathcal{O}(x^4) \\ u &= -10x_2 - 200x_1^2 \end{aligned} \quad (3.97)$$

while for  $\lambda = -\frac{1}{2}$ , the solution is

$$a = 2.732, b = \sqrt{2}, c = 8.90 \quad (3.98)$$

giving

$$\begin{aligned} V &= 2.732x_1^2 + \sqrt{2}x_2^2 + 8.9x_1^2x_2 \\ \mathcal{H}_* &= \mathcal{O}(x^4) \\ u &= -\sqrt{2}x_2 - 4.45x_1^2 \end{aligned} \tag{3.99}$$

For both cases, we ensure  $\dot{V}$  is zero through third order, and in fact, the fourth-order term that makes  $\dot{V}$  positive is  $2cx_1^2x_2^2$ .

### 3.5 SDRE Techniques

Herein we apply the SDRE nonlinear regulator theory to (3.1) in Section 3.5.1, and the SDRE nonlinear  $H_\infty$  control theory to (3.1) in Section 3.5.2.

#### 3.5.1 Nonlinear Regulation Via the SDRE Method

For this example we shall choose to minimize the infinite-horizon cost function

$$\text{minimize } J = \frac{1}{2} \int_{t_0}^{\infty} [x^T Q(x)x + u^T R(x)u] dt \tag{3.100}$$

with  $R(x) = \rho^2$  and

$$Q(x) = \begin{bmatrix} \mu^2 & 0 \\ 0 & 1 \end{bmatrix} = H^T(x)H(x) = \begin{bmatrix} \mu & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu & 0 \\ 0 & 1 \end{bmatrix}$$

so that  $\mu$  allows us to set a relative weight on state deviations from zero, and (3.100) is guaranteed convex since  $Q$  and  $R$  are positive definite. The scalar weights  $\rho$  and  $\mu$  will be treated as design parameters, the effects of which will be observed later.

Recalling that we need to factor the drift term in (3.1) as  $a(x) = A(x)x$ , we define

$$A_1(x) = \begin{bmatrix} x_2 & 0 \\ x_1 & 0 \end{bmatrix}, \quad A_2(x) = \begin{bmatrix} 0 & x_1 \\ x_1 & 0 \end{bmatrix} \tag{3.101}$$

and observe that (recalling Section 2.5.1)

$$A(\alpha, x) = \alpha A_1(x) + (1 - \alpha) A_2(x) = \begin{bmatrix} \alpha x_2 & (1 - \alpha) x_1 \\ x_1 & 0 \end{bmatrix} \quad (3.102)$$

is a valid SDC parametrization for all  $\alpha \in \mathcal{R}$ . In fact, it can easily be verified that (3.102) gives all valid SDC parametrizations of  $a(x)$ , and  $A_1$  corresponds to  $\alpha = 1$ , while  $A_2$  corresponds to  $\alpha = 0$ . Recalling that we desire controllable SDC parametrizations, we compute controllability matrices for  $A_1$ ,  $A_2$ , and  $A(\alpha, x)$  and respectively obtain

$$M_{c_1}(x) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad M_{c_2}(x) = \begin{bmatrix} 0 & x_1 \\ 1 & 0 \end{bmatrix}, \quad M_{c_\alpha}(x) = \begin{bmatrix} 0 & (1 - \alpha) x_1 \\ 1 & 0 \end{bmatrix} \quad (3.103)$$

so that  $A_1$  is an uncontrollable parametrization, and  $A_2$  and  $A(\alpha, x)$  are controllable so long as  $x_1 \neq 0$ . We note the similarity between the lack of existence of a controllable SDC parametrization for  $x_1 = 0$  to the lack of a well-defined relative degree for (3.1) at  $x_1 = 0$  observed in Section 3.2, and just as in that section, proceed with the SDRE nonlinear regulator algorithm in hopes of obtaining a stabilizing control in spite of technically failing to meet the controllability assumption. We also note that by our choice of  $H(x)$ , we have guaranteed observability of  $(H, A(\alpha, x))$  for any SDC parametrization  $A(\alpha, x)$ .

For simplicity, we will use the SDC parametrization  $A_2$  in the remaining development. With it and the choices for  $Q$  and  $R$  given above, and  $B = [0 \ 1]^T$ , the state feedback SDRE (2.142) becomes

$$\begin{aligned} & \begin{bmatrix} 0 & x_1 \\ x_1 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & x_1 \\ x_1 & 0 \end{bmatrix} \\ & - \frac{1}{\rho^2} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} \mu^2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned} \quad (3.104)$$

which simplifies to

$$x_1 \begin{bmatrix} 2p_{12} & p_{11} + p_{22} \\ p_{11} + p_{22} & 2p_{12} \end{bmatrix} - \frac{1}{\rho^2} \begin{bmatrix} p_{12}^2 & p_{12}p_{22} \\ p_{12}p_{22} & p_{22}^2 \end{bmatrix} + \begin{bmatrix} \mu^2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.105)$$

Equation (3.105) yields three nonlinear scalar equations

$$2x_1p_{12} - \frac{p_{12}^2}{\rho^2} + \mu^2 = 0 \quad (3.106)$$

$$2x_1p_{12} - \frac{p_{22}^2}{\rho^2} + 1 = 0 \quad (3.107)$$

$$(p_{11} + p_{22})x_1 - \frac{p_{12}p_{22}}{\rho^2} = 0 \quad (3.108)$$

only the first two of which need to be solved to obtain the control

$$u = -R^{-1}b^T Px = -\frac{1}{\rho^2}[0 \ 1] \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} x = -\frac{1}{\rho^2}(p_{12}x_1 + p_{22}x_2) \quad (3.109)$$

Equations (3.106) and (3.107) can be solved sequentially to give

$$p_{12}(x_1) = \rho^2 x_1 \pm \rho(\rho^2 x_1^2 + \mu^2)^{\frac{1}{2}} \quad (3.110)$$

$$p_{22}(x_1) = \rho(1 + 2x_1p_{12})^{\frac{1}{2}} \quad (3.111)$$

where we have taken the positive square root of  $p_{22}(x_1)$  in (3.111) to obtain a positive definite  $p$ , and have left the sign of the square root in (3.110) undetermined since we don't know which sign produces a stabilizing solution. To determine if a stabilizing solution to (3.104) exists, we expand  $p_{12}$  and  $p_{22}$  in a Taylor series about  $x_1 = 0$  (through first order) to obtain

$$p_{12}(x_1) = \pm\rho\mu + \rho^2 x_1 + \dots \quad (3.112)$$

$$p_{22}(x_1) = \rho + \rho(1 \pm \mu)x_1 + \dots \quad (3.113)$$

Using these expressions we obtain a second-order approximation for the control

$$u = \mp\frac{\mu}{\rho}x_1 - \frac{1}{\rho}x_2 - x_1^2 - \frac{1}{\rho}(1 \pm \mu)x_1x_2 \quad (3.114)$$

which, upon substitution into (3.1), gives the closed loop system equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_1x_2 \\ \mp\frac{\mu}{\rho}x_1 - \frac{1}{\rho}x_2 - \frac{1}{\rho}(1 \pm \mu)x_1x_2 \end{bmatrix} \quad (3.115)$$

which are in a form whose stability we can analyze via the Center Manifold Theorem (Theorem 2.1.7). Identifying the coordinate transformation

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \mu & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (3.116)$$

(3.115) can be transformed to

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\rho} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} h_1(z) \\ \mu z_2^2 + z_1 z_2 \end{bmatrix} \quad (3.117)$$

where  $h_1$ , a complicated function of the design weights and transformed state  $z$ , is irrelevant to stability analysis. Clearly, the  $\dot{z}_1$  equation is LAS for  $\rho > 0$ , while from Theorem 2.1.7 we must have  $z_1 = \pi(z_2) = c_1 z_2^2 + c_2 z_2^3 + \dots$  on the center manifold. Thus, the stability of the  $\dot{z}_2$  equation is determined by

$$\dot{z}_2 = \mu z_2^2 + z_2(c_1 z_2^2 + c_2 z_2^3 + \dots) = \mu z_2^2 + \mathcal{O}(z_2^3) \quad (3.118)$$

which from Lemmas 2.1.2 and 2.1.3 we may conclude to be unstable for any nonzero value of  $\mu$ . Therefore, there is no stabilizing solution to (3.104), which may be attributed to the fact that we are penalizing linear deviations of  $x_1$  from zero, yet the  $x_1$  state is not linearly stabilizable at  $x_1 = 0$ . This leads us to consider the case where  $\mu = 0$ , yielding solutions of (3.106)

$$p_{12}(x_1) = 0, \quad 2\rho^2 x_1 \quad (3.119)$$

and corresponding solutions of (3.107)

$$p_{22}(x_1) = \rho, \quad \rho(1 + 4\rho^2 x_1^2)^{\frac{1}{2}} \quad (3.120)$$

The first set of solutions to (3.119) and (3.120) yield the control

$$u = -\frac{1}{\rho} x_2$$

which is clearly not stabilizing since it is linear. The second expression for  $p_{22}$  in (3.120) can be expanded in a Taylor series about  $x_1 = 0$  to second order (the second-order coefficient is zero)

$$p_{22}(x_1) = \rho \quad (3.121)$$

so that to second order the control becomes

$$u = -\frac{1}{\rho}x_2 - 2x_1^2 \quad (3.122)$$

The control (3.122) gives a resulting closed loop system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{\rho} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} x_1x_2 \\ -x_1^2 \end{bmatrix} \quad (3.123)$$

Again using center manifold theory, it can be determined that the  $\dot{x}_1$  equation becomes

$$\dot{x}_1 = -\rho x_1^3 + \mathcal{O}(x_1^4) \quad (3.124)$$

which from Lemmas 2.1.2 and 2.1.3 we may conclude to be locally stable for any positive value of  $\rho$ . Thus, the SDRE nonlinear regulator does indeed provide a locally stabilizing control for (3.1), provided that our linearly uncontrollable state is also linearly unobservable in the cost function. This phenomenon of the SDRE method stabilizing uncontrollable (in the sense of not having a controllable SDC parametrization) modes is very interesting, and was also observed by Parrish in [53].

We note that in the final control (3.122), the gain on the controllable  $x_2$  state increases directly in proportion to a decrease in the penalty on control usage,  $\rho$ , just as we might expect from LQR theory. Also, if the state weighting on  $x_2$  is changed from 1 to  $k^2$ , the control becomes

$$u = -\frac{k}{\rho}x_2 - 2x_1^2 \quad (3.125)$$

so that now the gain on the controllable state is determined by the ratio  $\frac{k}{\rho}$ , and control usage will increase proportionately with increasing penalty on state deviations.

### 3.5.2 Nonlinear $H_\infty$ Control Via the SDRE Method

We now proceed to apply the full state feedback nonlinear  $H_\infty$  control theory of Section 2.5.2 to (3.1). We first need to establish that (3.1) meets the required assumptions. Recalling our convention for the system equations

$$\dot{x} = a(x) + b(x)u + g(x)d$$

$$z = h(x) + d_{12}(x)u$$

$$d_{12}^T(x)[h(x) \ d_{12}(x)] = [0 \ I] \quad (3.126)$$

we identify  $a$ ,  $b$ , and  $g$  as before and  $h(x) = [x_2 - x_1^2 \ 0]^T$ ,  $d_{12} = [0 \ 1]^T$ , so that indeed  $a(0) = 0$ ,  $h(0) = 0$ , and the last equation in (3.126) is satisfied. Next, we establish our SDC parametrizations of (3.126) as

$$\begin{aligned} a(x) &= A_2(x)x \\ h(x) &= H(x)x \end{aligned} \quad (3.127)$$

where

$$A_2(x) = \begin{bmatrix} 0 & x_1 \\ x_1 & 0 \end{bmatrix} \quad (3.128)$$

as in Section 3.5.1, and we take

$$H(x) = \begin{bmatrix} -x_1 & 1 \\ 0 & 0 \end{bmatrix} \quad (3.129)$$

noting that (3.129) is the unique SDC parametrization for  $h$ . With these choices of  $A_2(x)$  and  $H(x)$ , we find that

$$A_2(0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.130)$$

$$H(0) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (3.131)$$

giving an observability matrix

$$M_o = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.132)$$

which is clearly rank one. Since both eigenvalues of  $A_2(0)$  are zero, (3.132) allows us to conclude that the linearization of (3.126) is not detectable, and thus the SDRE nonlinear  $H_\infty$  solution method

is not guaranteed to be LAS for this system by the standard theory involving LAS linearizations. Since the SDRE must first be solved for either of the methods of Section 2.5.2, we first attempt the second approach. Then we attempt to apply the first technique, to see if it works and to give insight into reasons for failure if that is the end result.

We need to solve the state feedback nonlinear  $H_\infty$  SDRE

$$A^T(x)P + PA(x) - P[B(x)B^T(x) - \frac{1}{\gamma^2}G(x)G^T(x)]P + H^T(x)H(x) = 0 \quad (3.133)$$

which becomes after substitution of the proper values

$$\begin{aligned} & \begin{bmatrix} 0 & x_1 \\ x_1 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & x_1 \\ x_1 & 0 \end{bmatrix} - \left(\frac{1}{\gamma^2} - 1\right) \\ & \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} -x_1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -x_1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned} \quad (3.134)$$

Again defining  $\lambda = \left(\frac{1}{\gamma^2} - 1\right)$ , (3.134) simplifies to

$$x_1 \begin{bmatrix} 2p_{12} & p_{11} + p_{22} \\ p_{11} + p_{22} & 2p_{12} \end{bmatrix} + \lambda \begin{bmatrix} p_{12}^2 & p_{12}p_{22} \\ p_{12}p_{22} & p_{22}^2 \end{bmatrix} + \begin{bmatrix} -x_1^2 & -x_1 \\ -x_1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.135)$$

Equation (3.135) yields three nonlinear scalar equations

$$2x_1p_{12} + \lambda p_{12}^2 + x_1^2 = 0 \quad (3.136)$$

$$2x_1p_{12} + \lambda p_{22}^2 + 1 = 0 \quad (3.137)$$

$$(p_{11} + p_{22})x_1 + \lambda p_{12}p_{22} - x_1 = 0 \quad (3.138)$$

only the first two of which need to be solved to obtain the control (2.171)

$$u = -b^T P x = -[0 \ 1] \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} x = -p_{12}x_1 - p_{22}x_2 \quad (3.139)$$

Equations (3.136) and (3.137) can sequentially be solved analytically to give

$$p_{12}(x_1) = \frac{x_1}{\lambda}(-1 \pm \sqrt{1 - \lambda}) \quad (3.140)$$

$$p_{22}(x_1) = \left[ \frac{-1}{\lambda} + \frac{2x_1^2}{\lambda^2}(1 \mp \sqrt{1 - \lambda}) \right]^{\frac{1}{2}} \quad (3.141)$$

where we have taken the positive square root to obtain (3.141), and have left the sign of the square root as yet undetermined in (3.140). This sign will be chosen to yield a stabilizing solution. We note here that (3.136) has been solved exactly, but although (3.137) appears to also have been solved exactly, it actually has not. This is because  $p_{12}$  is linear in  $x_1$ , so that if we expand  $p_{22} = p_0 + p_1 x_1 + p_2 x_1^2 + \dots$ , the first-order equation in  $x_1$  resulting from (3.137) implies  $p_1 = 0$ . Thus there is no way to satisfy the second-order equation in  $x_1$ . To determine if a stabilizing solution to (3.134) exists, we use the Taylor series expansion of  $p_{22}$  about  $x_1 = 0$  to first order to obtain

$$p_{22}(x_1) = \sqrt{\frac{-1}{\lambda}} \quad (3.142)$$

We then define  $k_1 = -1 - \sqrt{1 - \lambda}$ ,  $k_2 = -1 + \sqrt{1 - \lambda}$ , and use these expressions to get a second-order approximation for the control

$$u = -\frac{k_i}{\lambda} x_1^2 - \sqrt{-\frac{1}{\lambda}} x_2 \quad (3.143)$$

where we take  $i = 1$  or  $2$  depending on which sign of the square root term in  $p_{12}$  we wish to use. If we recall all other stabilizing controls we have obtained in this chapter, we observe that each one was of the general form  $u = ax_1^2 + bx_2$ , where  $a$  and  $b$  were both negative. In fact, we know  $b$  must be negative, and it is easily proven using center manifold theory that we must have  $a < -1$  to obtain a stable closed loop system. Using this insight, we should expect that using  $k_1$  in (3.143) would give the stabilizing solution, and we therefore proceed under this assumption. For  $\lambda = -\frac{1}{100}$ , we obtain the control

$$u = -200x_1^2 - 10x_2 \quad (3.144)$$

whereas for  $\lambda = -\frac{1}{2}$ , we obtain

$$u = -4.45x_1^2 - \sqrt{2}x_2 \quad (3.145)$$

and recalling (3.97) and (3.99), we see that the above matches our earlier solutions from the Hamilton-Jacobi-Isaacs equations exactly. Now, even though we know (3.144) and (3.145) are stabilizing from other sources, we would like to know if this technique provides us a locally positive

definite Lyapunov function to establish closed loop stability on its own. Recalling that we have

$$V_x(x) = 2x^T P(x) = \left[ 2x_1 p_{11} + \frac{2k_1}{\lambda} x_1 x_2 \quad \frac{2k_1}{\lambda} x_1^2 + 2\sqrt{\frac{-1}{\lambda}} x_2 - \frac{2k_1}{\lambda^2} \sqrt{\frac{-1}{\lambda}} x_1^2 x_2 \right] \quad (3.146)$$

we need to solve simultaneously the integral equations

$$V = \int \left( 2x_1 p_{11} + \frac{2k_1}{\lambda} x_1 x_2 \right) dx_1 \quad (3.147)$$

$$= \int \left( \frac{2k_1}{\lambda} x_1^2 + 2\sqrt{\frac{-1}{\lambda}} x_2 - \frac{2k_1}{\lambda^2} \sqrt{\frac{-1}{\lambda}} x_1^2 x_2 \right) dx_2 \quad (3.148)$$

Equation (3.147) can be solved to give

$$V(x) = p_{11} x_1^2 + \frac{k_1}{\lambda} x_1^2 x_2 + f(x_2) \quad (3.149)$$

and (3.148) can be solved to give

$$V(x) = \frac{2k_1}{\lambda} x_1^2 x_2 + \sqrt{\frac{-1}{\lambda}} x_2^2 + g(x_1) \quad (3.150)$$

so that we must have  $f(x_2) = \sqrt{\frac{-1}{\lambda}} x_2^2$  and  $g(x_1) = p_{11} x_1^2$ . However, there is still a factor of two discrepancy between the  $x_1^2 x_2$  coefficients in the above two expressions for  $V$ . Now, recall that we have assumed  $P(x)$  to be symmetric, but there is really no need to enforce this assumption if we are solving the SDRE analytically. Careful examination of the Hamilton-Jacobi-Isacs equation for this problem with this symmetry assumption removed leads to three scalar equations. Two of these equations are the same as (3.136) and (3.138), while (3.137) becomes

$$2x_1 p_{21} + \lambda p_{22}^2 + 1 = 0 \quad (3.151)$$

where the only change is that  $p_{12}$  has become  $p_{21}$  in (3.151). Thus, there is no symmetry requirement on  $P(x)$ , and, in fact, to assure that  $V_x(x) = 2x^T P(x)$  has a consistent solution  $V$ , we can set  $p_{21} = 2p_{12}$ . We note this does not affect the optimal control derived previously, since for nonsymmetric  $P$ , the optimal control is in fact given by

$$u_* = -b^T P^T x = -[p_{12} \ p_{22}]x$$

and the first-order solution for  $p_{22}$  does not change under this modification. Making this change we obtain the Lyapunov function

$$V(x) = p_{11}x_1^2 + \frac{2k_1}{\lambda}x_1^2x_2 + \sqrt{\frac{-1}{\lambda}}x_2^2 \quad (3.152)$$

which does indeed solve the required set of partial differential equations. Now, (3.138) can be solved to first order to give  $p_{11} = 1 + \sqrt{1 - \frac{1}{\lambda}}$  and thus, since  $\lambda < 0$ , (3.152) is locally positive definite, and we can indeed conclude that the origin in the closed loop system is LAS. In fact, when the appropriate values of  $\lambda$  are substituted, (3.152) agrees exactly with the Lyapunov functions found in Section 3.4, by making the associations  $a = p_{11}$ ,  $b = \sqrt{\frac{-1}{\lambda}}$ , and  $c = \frac{2k_1}{\lambda}$ . We also point out that  $V$  in (3.152) can be written in the form  $V = x^T W(x)x$ , where  $W(0)$  is a positive definite matrix, and the off-diagonal elements of  $W(x)$  are nonunique, so that  $W$  may or may not be symmetric.

Now, to apply the first approach to SDRE nonlinear  $H_\infty$  control, we also need to compute the contribution to the control from the  $M_P$  term we previously defined as

$$M_P = x^T \frac{\partial P}{\partial x} \quad (3.153)$$

which is given by  $u_M = -\frac{1}{2}b^T M_P^T x$ . Using the product rule  $\frac{\partial}{\partial x}(Px) = P + x^T \frac{\partial P}{\partial x} = P + M_P$ , an expression for the matrix function  $M_P$  can be derived without having to invoke tensor operations. The resulting expression for the  $ij$ th entry of  $M_P$  is

$$M_{Pij} = \sum_{k=1}^n x_k \frac{\partial P_{ik}}{\partial x_j} \quad (3.154)$$

Using the symmetric first-order solution to the  $H_\infty$  SDRE, (3.133),

$$P = \begin{bmatrix} 1 + \sqrt{1 - \frac{1}{\lambda}} & \frac{k_1}{\lambda}x_1 \\ \frac{k_1}{\lambda}x_1 & \sqrt{\frac{-1}{\lambda}} \end{bmatrix}$$

$M_P$  can be computed as

$$M_P = \begin{bmatrix} \frac{k_1}{\lambda}x_2 & 0 \\ \frac{k_1}{\lambda}x_1 & 0 \end{bmatrix} \quad (3.155)$$

so that  $u_M = 0$ , producing the same control as in both previous methods. Also, we find that the second bracketed term in (2.173) reduces to  $2\frac{k_1}{\lambda}x_1^2x_2^2$ , while our inability to solve (3.137) exactly also

leaves a term equal to  $2\frac{k_1}{\lambda}x_1^2x_2^2$ . Thus, our solution to the Hamilton-Jacobi-Isaacs equation using this approach is again zero to third order, with the next remaining term equal to  $4\frac{k_1}{\lambda}x_1^2x_2^2 = 2cx_1^2x_2^2$ .

We obtain our Lyapunov function from this approach via

$$V(x) = x^T P(x)x = x^T \begin{bmatrix} p_{11} & \frac{k_1}{\lambda}x_1 \\ \frac{k_1}{\lambda}x_1 & \sqrt{\frac{-1}{\lambda}} \end{bmatrix} x = p_{11}x_1^2 + \frac{2k_1}{\lambda}x_1^2x_2 + \sqrt{\frac{-1}{\lambda}}x_2^2 \quad (3.156)$$

so that in fact we have obtained the same solution to the nonlinear local  $H_\infty$  control problem using all three approaches.

Summarizing, it appears that both SDRE approaches may be used to solve the local state feedback nonlinear  $H_\infty$  control problem analytically. Using the first approach of Section 2.5.2, we propose  $V = x^T P(x)x$ ,  $P = P^T$ , so that  $V_x = x^T(2P + M_P)$ , and we must solve the  $H_\infty$  SDRE (2.161) plus the additional inequality from the second bracketed term in (2.173). The advantages of this method are that the symmetry of  $P$  make the solution of (2.161) simpler, and that we are guaranteed the existence of a locally positive definite Lyapunov function of known form under certain known conditions. The main disadvantage is having to compute  $M_P$ , and dealing with the second bracketed term in (2.173). Using the second approach of Section 2.5.2, we propose  $V_x = 2x^T P$ , and we must solve the  $H_\infty$  SDRE (2.161), followed by the  $n$ -dimensional set of partial differential equations dictated by  $V_x$  for the Lyapunov function  $V$ . The advantage here is we only need to solve the state feedback  $H_\infty$  SDRE, and in the limit as  $\gamma$  approaches infinity, (2.161) approaches the appropriate Hamilton-Jacobi-Bellman optimal control equation. The main disadvantage is that we can't as yet guarantee the existence of a locally positive definite Lyapunov function, and we must solve  $n$  pde's to obtain a candidate. In this example problem we were able to find such a Lyapunov function successfully, but it required taking a nonsymmetric solution  $P_u$  to the state feedback  $H_\infty$  SDRE, (2.161). In this case, the solution was underdetermined, and  $p_{21}$  was chosen in an ad hoc manner, based on what we knew the answer should be to obtain a consistent Lyapunov function. Thus, it may not always be straightforward to construct the appropriate nonsymmetric SDRE solution. In doing so for this example problem, we obtained equal expressions for  $V_x$  and hence  $V$  using both methods. This implies  $x^T(2P_u) = x^T(2P + M_P)$ , and yet from simple calculation we find  $2P_u \neq 2P + M_P$ .

This observation suggests that determining more specific conditions under which both methods may be used to obtain the same result may be difficult to accomplish. Nevertheless, this topic is pursued in Section 4.5, and partial results are obtained relating the two solution approaches for the case of symmetric  $P$ .

A final comment on the SDRE nonlinear  $H_\infty$  control techniques is that we have not addressed the effects of different SDC parametrizations on solvability of (2.161) and (2.173). Using the uncontrollable SDC parametrization  $A_1(x)$ , it is straightforward to solve (2.161) and show that no stabilizing solution exists. Recalling,  $A(\alpha, x) = \alpha A_1(x) + (1 - \alpha)A_2(x)$ , we have thus solved the problem for  $\alpha = 1, 0$ , respectively corresponding to the solutions for  $A_1$  and  $A_2$ . For any value of  $\alpha$  other than 1 or 0, the three nonlinear scalar equations resulting from (2.161) become coupled, and cannot be solved sequentially. Thus, it is hard to determine analytically the effects of  $\alpha$  on solvability of the example problem. Using a different value of  $\alpha$  and hence a different SDC parametrization of  $a$ , it may be possible to solve (2.161) to higher than third order, perhaps even exactly. This research has not been carried out, however, as the more fruitful line of research of numerical solution approaches has been pursued.

In this chapter we have applied the nonlinear control design methods of Chapter 2 to a motivational, second-order example problem. In doing so, we have illustrated the principles, strengths, and weaknesses associated with each method, allowing us to see where the SDRE methods may offer advantages. The example problem also raises a number of questions about SDRE solution approaches, which will be addressed in Chapter 4.

## IV. Insights into Design Issues

In this chapter we address several issues concerning SDRE control algorithms that were introduced in Chapters 2 and 3. These issues are primarily related to two main thrusts: the relationship between the suboptimal SDRE regulator and the corresponding optimal controller for the infinite time problem, and solution approaches to HJIs via SDREs. With respect to the former, we show how convexity affects sufficient conditions for local and global optimality, we derive and interpret a simplified necessary condition for the SDRE method to give the optimal solution, and we develop a numerical algorithm to enforce satisfaction of the simplified necessary condition. In the latter regard we consider a number of varied topics: two proposed solution approaches, their Lyapunov functions and corresponding relationships, symmetry of SDRE solutions, numerical versus analytical SDRE solutions, necessary conditions for solvability, and strict inequality versus equation problem formulations.

### 4.1 Partial Derivatives of Vector Matrix Products with $x$ Dependency

Quite often in the course of derivations and proofs involved with SDRE control, it is necessary to compute partial derivatives with respect to  $x$  of vector matrix products having explicit  $x$  dependency (see Sections 2.5.1 and 2.5.2 for example). Since formulas for such are not readily available in texts, original derivations of the needed formulas are given in this section. Suppose we wish to compute  $\frac{\partial z^T P}{\partial x}$  where  $z(x) \in \mathcal{R}^n$  is a vector valued function of  $x$  and  $P(x) \in \mathcal{R}^{n \times n}$  is similarly a matrix valued function of  $x$ . If we let

$$m^T \equiv z^T P \tag{4.1}$$

then  $m^T(x) \in \mathcal{R}^n$  is a (row) vector function of  $x$  with elements

$$m_i = \sum_{k=1}^n z_k P_{ki} \tag{4.2}$$

Thus, using standard convention for the partial derivative of a row vector with respect to a vector, we have

$$\frac{\partial z^T P}{\partial x} = \frac{\partial m^T}{\partial x} \equiv Y = \begin{bmatrix} \frac{\partial m_1}{\partial x_1} & \frac{\partial m_2}{\partial x_1} & \dots & \frac{\partial m_n}{\partial x_1} \\ \frac{\partial m_1}{\partial x_2} & \frac{\partial m_2}{\partial x_2} & \dots & \frac{\partial m_n}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial m_1}{\partial x_n} & \frac{\partial m_2}{\partial x_n} & \dots & \frac{\partial m_n}{\partial x_n} \end{bmatrix} \quad (4.3)$$

and we see  $Y$  is a matrix function of  $x$  as expected. Using (4.2) and the repeated index convention to represent summation over the repeated index ( $a_k b_k = \sum_k a_k b_k$ ), we find

$$Y = \begin{bmatrix} \frac{\partial z_k}{\partial x_1} P_{k1} + z_k \frac{\partial P_{k1}}{\partial x_1} & \frac{\partial z_k}{\partial x_1} P_{k2} + z_k \frac{\partial P_{k2}}{\partial x_1} & \dots & \frac{\partial z_k}{\partial x_1} P_{kn} + z_k \frac{\partial P_{kn}}{\partial x_1} \\ \frac{\partial z_k}{\partial x_2} P_{k1} + z_k \frac{\partial P_{k1}}{\partial x_2} & \frac{\partial z_k}{\partial x_2} P_{k2} + z_k \frac{\partial P_{k2}}{\partial x_2} & \dots & \frac{\partial z_k}{\partial x_2} P_{kn} + z_k \frac{\partial P_{kn}}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z_k}{\partial x_n} P_{k1} + z_k \frac{\partial P_{k1}}{\partial x_n} & \frac{\partial z_k}{\partial x_n} P_{k2} + z_k \frac{\partial P_{k2}}{\partial x_n} & \dots & \frac{\partial z_k}{\partial x_n} P_{kn} + z_k \frac{\partial P_{kn}}{\partial x_n} \end{bmatrix} \quad (4.4)$$

so that the  $ij$ th element of  $Y$  can be written

$$Y_{ij} = \frac{\partial z_k}{\partial x_i} P_{kj} + z_k \frac{\partial P_{kj}}{\partial x_i} \quad (4.5)$$

Thus, by observing

$$\frac{\partial z^T}{\partial x} P = \begin{bmatrix} \frac{\partial z_k}{\partial x_1} P_{k1} & \frac{\partial z_k}{\partial x_1} P_{k2} & \dots & \frac{\partial z_k}{\partial x_1} P_{kn} \\ \frac{\partial z_k}{\partial x_2} P_{k1} & \frac{\partial z_k}{\partial x_2} P_{k2} & \dots & \frac{\partial z_k}{\partial x_2} P_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z_k}{\partial x_n} P_{k1} & \frac{\partial z_k}{\partial x_n} P_{k2} & \dots & \frac{\partial z_k}{\partial x_n} P_{kn} \end{bmatrix} \quad (4.6)$$

and adopting the convention

$$z^T \frac{\partial P}{\partial x} = \begin{bmatrix} z_k \frac{\partial P_{k1}}{\partial x_1} & z_k \frac{\partial P_{k2}}{\partial x_1} & \dots & z_k \frac{\partial P_{kn}}{\partial x_1} \\ z_k \frac{\partial P_{k1}}{\partial x_2} & z_k \frac{\partial P_{k2}}{\partial x_2} & \dots & z_k \frac{\partial P_{kn}}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ z_k \frac{\partial P_{k1}}{\partial x_n} & z_k \frac{\partial P_{k2}}{\partial x_n} & \dots & z_k \frac{\partial P_{kn}}{\partial x_n} \end{bmatrix} \quad (4.7)$$

it can be seen that

$$\frac{\partial z^T P}{\partial x} = \frac{\partial z^T}{\partial x} P + z^T \frac{\partial P}{\partial x} \quad (4.8)$$

In the notation in (4.7) we consider  $z^T \frac{\partial P}{\partial x}$  a single object, and note that the  $x_i$  with respect to which the partial is being taken remains the same in each row of the matrix. This corresponds to standard notation in that we are taking the partial of a row vector, as indicated by the premultiplying  $z^T$ . By letting  $z = x$ , it is clear from (4.8) that

$$\frac{\partial x^T P}{\partial x} = P + x^T \frac{\partial P}{\partial x} \quad (4.9)$$

By simply substituting  $P^T$  for  $P$  above and carrying out the same operations, we find

$$\frac{\partial z^T P^T}{\partial x} = \frac{\partial z^T}{\partial x} P^T + z^T \frac{\partial P^T}{\partial x} \quad (4.10)$$

which in the case of  $z = x$  simplifies to

$$\frac{\partial x^T P^T}{\partial x} = P^T + x^T \frac{\partial P^T}{\partial x} \quad (4.11)$$

We can perform similar operations when  $P$  is postmultiplied by  $z$  to obtain

$$\frac{\partial Pz}{\partial x} = P \frac{\partial z}{\partial x} + \frac{\partial P}{\partial x} z \quad (4.12)$$

where we have adopted the convention

$$\frac{\partial P}{\partial x} z = \begin{bmatrix} z_k \frac{\partial P_{1k}}{\partial x_1} & z_k \frac{\partial P_{1k}}{\partial x_2} & \cdots & z_k \frac{\partial P_{1k}}{\partial x_n} \\ z_k \frac{\partial P_{2k}}{\partial x_1} & z_k \frac{\partial P_{2k}}{\partial x_2} & \cdots & z_k \frac{\partial P_{2k}}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ z_k \frac{\partial P_{nk}}{\partial x_1} & z_k \frac{\partial P_{nk}}{\partial x_2} & \cdots & z_k \frac{\partial P_{nk}}{\partial x_n} \end{bmatrix} \quad (4.13)$$

When  $z = x$  (4.12) simplifies to

$$\frac{\partial Px}{\partial x} = P + \frac{\partial P}{\partial x} x \quad (4.14)$$

For completeness we also have

$$\frac{\partial P^T z}{\partial x} = P^T \frac{\partial z}{\partial x} + \frac{\partial P^T}{\partial x} z \quad (4.15)$$

which simplifies to

$$\frac{\partial P^T x}{\partial x} = P^T + \frac{\partial P^T}{\partial x} x \quad (4.16)$$

in the case where  $z = x$ . Note that in adopting these conventions, we have preserved the conventions that

$$\frac{\partial m^T}{\partial x} = \left( \frac{\partial m}{\partial x} \right)^T \quad (4.17)$$

and

$$(ab)^T = b^T a^T \quad (4.18)$$

Note that (4.18) even applies to the constructs in (4.13) and (4.7), so that for example

$$\left( z^T \frac{\partial P}{\partial x} \right)^T = \frac{\partial P^T}{\partial x} z \quad (4.19)$$

To conclude this section we relate the definitions given here to the notation of Sections 2.5.1 and 2.5.2. From (2.166) we see that  $M_P$  in Section 2.5.2 is given by (4.13), with the substitution  $z = x$ , while from (4.19) we see that  $M_P^T$  is given by (4.7) with the same substitution. Simple calculations verify that for any symmetric matrix  $P$

$$x^T M_P = M_P^T x = x^T \frac{\partial P}{\partial x} x \quad (4.20)$$

so that  $x^T P_x x$  in Section 2.5.1 may equivalently be replaced by any of the expressions in (4.20).

## 4.2 Convexity Concerns in SDRE Nonlinear Regulation

In considering existence and determination of optimal solutions to the constrained minimization problem of Section 2.5.1, the issue of convexity arises. For example, necessary and sufficient conditions for global optimality are most easily formulated for convex optimization problems [46]. Even in seeking local optima, convexity plays a role in determining sufficient conditions. In Sections 2.5.1 and 4.4, the first-order necessary criteria for optimal solutions to the constrained control problem are given. To obtain sufficient conditions for a strong local minimum, the additional Legendre-Clebsch, Weierstrauss, and Jacobi (nonconjugate point) conditions may be enforced, provided the system is normal [8]. It turns out that the Legendre-Clebsch and Weierstrauss conditions are related to convexity of the Hamiltonian

$$\mathcal{H} = x^T Q(x)x + u^T R(x)u + \lambda^T (a(x) + b(x)u) \quad (4.21)$$

with respect to the control  $u$ . These two conditions are satisfied if

$$\mathcal{H}_{uu} = R(x) > 0 \quad \forall x \in \mathcal{R}^n \quad (4.22)$$

which we always satisfy by assumption in our choice of  $R$ . Thus, if we choose  $R(x) > 0$  for all  $x$  and satisfy the SDRE Necessary Condition for Optimality, we need only satisfy the Jacobi and normality conditions to assure obtaining a locally minimizing control with respect to the cost function  $J$ .

On the other hand, global minimizations will rely on convexity of  $\mathcal{H}$  with respect to  $x$  and  $u$ . It is therefore of interest to know what kinds of state weighting matrix functions  $Q(x)$  yield

$$\frac{\partial^2 x^T Q x}{\partial x^2} > 0 \quad \forall x \quad (4.23)$$

Using the insight of (4.7) and (4.13), we have the following theorem, which provides a sufficient condition for convexity of the state component of  $\mathcal{H}$ .

**Theorem 4.2.1** *Consider the scalar function  $l = x^T Q(x)x$ . Choose  $Q(x) = Q_0 + Q_1(x)$  where  $Q_0$  is any symmetric positive definite matrix and  $Q_1(x)$  is equal to a diagonal matrix function  $Q_1(x) = \text{diag}(q_1(x_1), \dots, q_n(x_n))$ , where each  $q_i$  takes the form  $q_i(x_i) = c_{i0} + c_{i2}x_i^2 + c_{i4}x_i^4 + \dots + c_{is_i}x_i^{s_i}$ , with  $c_{ij} \geq 0$ ,  $j = 0, 2, 4, \dots, s_i$ . Then  $l$  is globally convex with respect to  $x$ .*

**Proof:** Straightforward application of either (4.8) and (4.7) or (4.12) and (4.13), followed by (4.20). ■

Note that Theorem 4.2.1 justifies the logical choice for state-dependent penalty weightings on the states which yield only sums of terms consisting of nonnegative coefficients multiplying higher order, even powers of individual states in the cost function. Thus, if  $Q(x)$  is chosen as per Theorem 4.2.1, each state  $x_i$  has a term in the cost function of the form  $c_{i0}x_i^2 + c_{i2}x_i^4 + c_{i4}x_i^6 + \dots + c_{is_i}x_i^{s_i+2}$ , where all the  $c_{ij}$  are greater than or equal to zero. Such a state weighting provides a steeper penalty for nonzero state deviations far from the origin than purely constant state weightings does, and thus can be expected to increase control gains far from the origin. We close this section by stating that most state weighting matrix function examples in this dissertation are chosen to satisfy the requirements

of Theorem 4.2.1 so as to obtain the desirable properties of convex functionals. In some cases we choose  $Q(x)$  to satisfy  $l_{xx} \geq 0$ , so that we have convexity on a reduced part of the state space.

### 4.3 Symmetry of $P$ in Proposed HJI Solution $V = x^T P x$

Recall now that in Section 2.5.2 two approaches to solving the HJI equation (SDRE) associated with the nonlinear  $H_\infty$  suboptimal control problem were proposed. One was to let  $V = x^T P x$ , with  $P = P^T$ , and attempt to solve the  $H_\infty$  SDRE (2.161) and additional inequality  $N(x) \leq 0$ . The other was to let  $V_x = 2x^T P$ , with  $P$  not necessarily symmetric, solve the SDRE, and attempt to solve the PDE above to obtain a (locally) positive definite  $V$ . We now consider a third alternative, motivated by the successful application of Method 2 with nonsymmetric  $P$  to the example problem of Chapter 3. This alternative is to consider a solution of the form  $V = x^T P x$ , with  $P$  not necessarily symmetric, and investigate the utility of removing the symmetry assumption on  $P$ . It turns out that no utility is added by removing the symmetry assumption on  $P$  since only the symmetric part contributes to the solution, as shown below.

Recall that we wish to solve the HJI

$$H^* = V_x f + z^T z - \gamma^2 d^T d = V_x \left[ Ax - \frac{1}{4} B B^T V_x^T + \frac{1}{4} G G^T V_x^T \right] + x^T H^T H x \leq 0 \quad (4.24)$$

With  $V = x^T P x$  and using (4.14) we have

$$V_x = x^T P^T + x^T \left[ P + \frac{\partial P}{\partial x} x \right] = x^T \left[ P + P^T + \frac{\partial P}{\partial x} x \right] \quad (4.25)$$

so that using (2.166) we see that (4.25) can be written

$$V_x = x^T [P + P^T + M_P] \quad (4.26)$$

Now, if we write  $P$  as the sum of its symmetric and skew-symmetric parts [42]

$$P = P_{sym} + P_{sk} \quad (4.27)$$

where

$$P_{sym} = \frac{1}{2}(P + P^T) \quad (4.28)$$

and

$$P_{sk} = \frac{1}{2}(P - P^T) \quad (4.29)$$

we immediately see from (4.28) that  $P + P^T = 2P_{sym}$ , so that (4.26) becomes

$$V_x = x^T [2P_{sym} + M_P] \quad (4.30)$$

Substituting (4.30) into (4.24) and simplifying we obtain

$$x^T [A^T P_{sym} + P_{sym} A + P_{sym} K P_{sym} + H^T H] x + x^T \left[ M_P (A + K P_{sym} + \frac{1}{4} K M_P^T) \right] x \leq 0 \quad (4.31)$$

where we have defined

$$K \equiv \frac{1}{\gamma^2} G G^T - B B^T \quad (4.32)$$

Thus, we can see that the SDRE part of (4.31) involves only the symmetric part of  $P$ , with the skew-symmetric part contributing nothing. Now we must consider the second part of (4.31), which contains the term  $M_P$ . If we decompose  $P$  according to (4.27) in (2.166), we can write  $M_P$  as  $M_P = M_1 + M_2$  where

$$M_1 \equiv \frac{\partial P_{sym}}{\partial x} x \quad (4.33)$$

$$M_2 \equiv \frac{\partial P_{sk}}{\partial x} x \quad (4.34)$$

Now, using the fact that for any skew-symmetric matrix  $P_{sk}^T = -P_{sk}$ , so that  $x^T P_{sk} x = 0$  for all  $x$ , we have

$$\frac{\partial x^T P_{sk} x}{\partial x} = x^T \left[ P_{sk}^T + P_{sk} + \frac{\partial P_{sk}}{\partial x} x \right] = 0 \quad (4.35)$$

so that

$$x^T \frac{\partial P_{sk}}{\partial x} x = x^T M_2 = 0 \quad (4.36)$$

Thus, (4.31) reduces to

$$x^T [A^T P_{sym} + P_{sym} A + P_{sym} K P_{sym} + H^T H] x + x^T \left[ M_1 (A + K P_{sym} + \frac{1}{4} K M_1^T) \right] x \leq 0 \quad (4.37)$$

so that indeed the solution to (4.24) depends only on the symmetric part of  $P$ .

#### 4.4 Simplification of SDRE Necessary Condition for Optimality

Recall that in the SDRE nonlinear regulator problem we wish to find the control  $u$  which solves the following optimal control problem equation (note the  $x$  dependency notation has been dropped for brevity):

$$0 = \min_u \mathcal{H} = \min_u \frac{1}{2} x^T Q x + \frac{1}{2} u^T R u + \lambda^T [A x + B u] \quad (4.38)$$

By invoking the first-order necessary condition  $\mathcal{H}_u = 0^T$  and assuming  $R = R^T$ , we obtain

$$u_o^T R + \lambda^T B = 0 \quad (4.39)$$

where we have used the notation  $u_o$  for the optimal control. Making the association  $\lambda = P x$  and solving we get

$$u_o(x) = -R^{-1} B^T P x \quad (4.40)$$

Substituting (4.40) into (4.38) and simplifying leads to

$$x^T [A^T P + P^T A - P^T B R^{-1} B^T P + Q] x = 0 \quad (4.41)$$

which can be solved by setting the term in brackets equal to zero, yielding the state-dependent Riccati equation in the case where we assume  $P = P^T$  as in Section 2.5.1. In [13] the authors give an additional vector equation of dimension  $n$  which must be satisfied in order for the SDRE method to satisfy the necessary conditions of the optimal nonlinear regulator problem. This equation, called the *SDRE Necessary Condition for Optimality*, is related to requiring the solution to follow the optimal costate vector trajectory. The equation given in [13] and repeated in Section 2.5.1 is (again assuming  $P = P^T$ )

$$\dot{P} x + \frac{1}{2} x^T \left[ \frac{\partial Q}{\partial x} x \right] + \frac{1}{2} x^T P B R^{-1} \left[ \frac{\partial R}{\partial x} x \right] R^{-1} B^T P x + x^T \left[ \frac{\partial A}{\partial x} x \right] P x - x^T P B R^{-1} \left[ \frac{\partial B^T}{\partial x} x \right] P x = 0 \quad (4.42)$$

which is a complicated equation involving third-order tensors and the time derivative of the  $P$  matrix. If we were to factor out an  $x$  to the right in (4.42) and then set the premultiplying term equal to zero, we would have  $n^2$  constraint equations to solve. In this section we derive an alternative form for (4.42) which is significantly simpler, also providing a physical explanation for the resulting

requirement and showing that for symmetric  $P$  the necessary condition for optimality actually places  $n(n-1)/2$  as opposed to  $n^2$  additional constraints on the control problem. The derivation of this simplified necessary condition takes advantage of the fact that (4.41) is satisfied by our choice of  $A$  and  $P$ , and is presented below.

We start by considering the most general case in which  $P$  is not necessarily symmetric. Differentiating  $\lambda = Px$ , for the closed loop system we obtain

$$\dot{\lambda} = \frac{\partial Px}{\partial x} \dot{x} = \left[ \frac{\partial P}{\partial x} x + P \right] (Ax + Bu_o) \quad (4.43)$$

Transposing, we get

$$\dot{\lambda}^T = (x^T A^T + u_o^T B^T) \left[ P^T + x^T \frac{\partial P^T}{\partial x} \right] \quad (4.44)$$

For optimality, we need to satisfy

$$\dot{\lambda}^T = -\frac{\partial \mathcal{H}}{\partial x} \quad (4.45)$$

Evaluating the right hand side of (4.45) (recalling that  $u$  and  $\lambda$  are considered variables independent of  $x$  for purposes of this operation) we find

$$\dot{\lambda}^T = -x^T Q - \frac{1}{2} x^T \left[ \frac{\partial Q}{\partial x} x \right] - \frac{1}{2} u_o^T \left[ \frac{\partial R}{\partial x} u_o \right] - \lambda^T \left[ \frac{\partial A}{\partial x} x + A + \frac{\partial B}{\partial x} u_o \right] \quad (4.46)$$

Using the fact that on the optimal trajectory  $u = u_o$  we have

$$\dot{\lambda}^T = -x^T Q - \frac{1}{2} x^T \left[ \frac{\partial Q}{\partial x} x \right] - \frac{1}{2} u_o^T \left[ \frac{\partial R}{\partial x} u_o \right] - \lambda^T \left[ \frac{\partial A}{\partial x} x + A + \frac{\partial B}{\partial x} u_o \right] \quad (4.47)$$

Now, recall that in satisfying (4.41) we are actually just satisfying

$$\mathcal{H}(x, u_o) = \frac{1}{2} x^T Q x + \frac{1}{2} u_o^T R u_o + x^T P^T [Ax + Bu_o] = 0 \quad (4.48)$$

for all  $x$ . Thus, taking the partial derivative of (4.48) with respect to  $x$  we get

$$\begin{aligned} & x^T Q + \frac{1}{2} x^T \left[ \frac{\partial Q}{\partial x} x \right] + \frac{1}{2} u_o^T \left[ \frac{\partial R}{\partial x} u_o \right] + [u_o^T R + x^T P^T B] \frac{\partial u_o}{\partial x} \\ & + x^T P^T \left[ \frac{\partial A}{\partial x} x + A + \frac{\partial B}{\partial x} u_o \right] + (x^T A^T + u_o^T B^T) \left[ P + \frac{\partial P}{\partial x} x \right] = 0 \end{aligned} \quad (4.49)$$

Replacing  $x^T P^T$  with  $\lambda^T$ , using (4.39) so that the coefficient of  $\frac{\partial u_o}{\partial x}$  vanishes, and rearranging gives

$$-x^T Q - \frac{1}{2} x^T \left[ \frac{\partial Q}{\partial x} x \right] - \frac{1}{2} u_o^T \left[ \frac{\partial R}{\partial x} u_o \right] - \lambda^T \left[ \frac{\partial A}{\partial x} x + A + \frac{\partial B}{\partial x} u_o \right] = (x^T A^T + u_o^T B^T) \left[ P + \frac{\partial P}{\partial x} x \right] \quad (4.50)$$

Substituting (4.50) into (4.47), we find

$$\dot{\lambda}^T = (x^T A^T + u_o^T B^T) \left[ P + \frac{\partial P}{\partial x} x \right] \quad (4.51)$$

Thus, comparing (4.51) to (4.44), we see that to achieve the optimal solution we must have

$$(x^T A^T + u_o^T B^T) \left[ P^T + x^T \frac{\partial P^T}{\partial x} \right] = (x^T A^T + u_o^T B^T) \left[ P + \frac{\partial P}{\partial x} x \right] \quad (4.52)$$

This is the simplified necessary condition for optimality which is significantly easier to work with than (4.42), for clearly a sufficient condition for (4.52) to hold is to have

$$P^T + x^T \frac{\partial P^T}{\partial x} = P + \frac{\partial P}{\partial x} x \quad (4.53)$$

If we now enforce the assumption that  $P = P^T$  as in [13], the SDRE simplifies to

$$A^T P + P A - P B R^{-1} B^T P + Q = 0 \quad (4.54)$$

while (4.53) becomes

$$x^T \frac{\partial P}{\partial x} = \frac{\partial P}{\partial x} x \quad (4.55)$$

or

$$x^T \frac{\partial P}{\partial x} - \frac{\partial P}{\partial x} x = 0 \quad (4.56)$$

An interpretation of (4.53) is provided by making use of an insight from Hamilton-Jacobi-Bellman theory. In HJB theory, the optimal costate differential equation is automatically satisfied when the HJB equation is solved (see [8]). The key reason is that  $\dot{\lambda}^T = V_x$ , where  $V$  is the optimal cost for the control problem. Thus, the costate vector is a perfect differential, so that its partial derivative with respect to  $x$  yields the Hessian matrix of  $V$ , which is symmetric. This symmetry is key to cancellation of all terms in the proof of satisfaction of  $\dot{\lambda}^T = -\mathcal{H}_x$  in [8]. This suggests (4.53) is equivalent to the requirement that the costate vector be a perfect differential, so that the optimal cost  $V$  exists and is well-defined. This conjecture is validated when one sets  $\frac{\partial \lambda}{\partial x} = \frac{\partial \lambda^T}{\partial x}$ , corresponding to the symmetric Hessian requirement for  $V$ . In so doing one obtains (4.53). Thus, satisfaction of (4.53) is in fact equivalent to existence of an optimal cost  $V$ , for which  $\lambda^T = x^T P^T$  is its differential (there is an if

and only if correspondence). This equivalence is stated in [45], and, in addition, a formula for an appropriate  $V$  (appropriate in the sense that  $V(0) = 0$ ) is given as

$$V(x) = 2x^T \int_0^1 P(tx)(tx)dt \quad (4.57)$$

Also, it is stated that if  $P$  is a positive definite matrix-valued function, then  $V$  is positive definite. Thus, (4.52) is required to be satisfied in order to achieve optimality, but satisfaction of (4.53) is at least sufficient (and perhaps necessary) for satisfying (4.52), and necessary in order to have a well defined ( $C^1$ ) optimal cost function exist. We additionally would like to have a positive definite solution  $P$  so that  $V$  is positive definite, enabling its use as a Lyapunov function for the system.

We see that in general (4.53) is a matrix equation in  $x$  which imposes  $n^2$  scalar constraints on the optimization problem. For  $P = P^T$  however, by examining the expanded version of (4.56) using (4.7) and (4.13), it can be seen that the diagonal elements of the left hand side of (4.56) are automatically zero due to the symmetry assumption on  $P$ . Also, the off-diagonal elements of the left hand side of (4.56) are equal to their symmetric counterparts due to the symmetry assumption. Thus, in the case of symmetric  $P$ , the left hand side of (4.56) is a symmetric matrix function so that (4.56) imposes  $k = n(n - 1)/2$  constraints on the problem.

#### 4.5 Relationships Between Solution Methods for SDRE Nonlinear $H_\infty$ Control

In this section we build on the insights of Sections 4.3 and 4.4 to develop relationships between Methods 1 and 2 proposed in Section 2.5.2 for solving nonlinear  $H_\infty$  control problems via the SDRE technique. We assume the existence of stabilizable and detectable factorizations. Suppose we solve the  $H_\infty$  SDRE

$$A^T P + PA + PKP + H^T H = 0 \quad (4.58)$$

where  $K$  is as defined in (4.32), and  $P$  is assumed symmetric. Then suppose we seek to use Method 2 of Section 2.5.2 so that we have  $V_x = 2x^T P$ , and from Section 4.4 we know that a solution  $V$  to the PDE above only exists if  $x^T P$  is a perfect differential of a scalar function of  $x$ , or equivalently if

$$M_P = M_P^T \quad (4.59)$$

where  $M_P$  is defined as in (2.166). If we suppose (4.59) holds, then  $V$  is given as in Section 4.4 by

$$V(x) = 2x^T \int_0^1 P(tx)(tx)dt \quad (4.60)$$

Now let us compare these results to the results of using Method 1 of Section 2.5.2. Using Method 1, suppose we also solve (4.58) and satisfy (4.59), and additionally our solution  $P$  is such that we have

$$N(x) = M_P A + \frac{1}{4} M_P K M_P + M_P K P = 0 \quad (4.61)$$

It is easily shown that simultaneously satisfying (4.58), (4.59), and (4.61) is the same as satisfying the algebraic Riccati equation

$$A^T P_1 + P_1 A + P_1 K P_1 + H^T H = 0 \quad (4.62)$$

where

$$P_1 = P + \frac{1}{2} M_P \quad (4.63)$$

However, the two Riccati equations (4.58) and (4.62) are the same, so that by uniqueness of solutions we must have  $P_1 = P$ , and thus by (4.63) we must have  $M_P = 0$ . This is seen alternatively by rearranging (4.61) to the form of an algebraic Riccati equation

$$M_P(A + KP) + (A + KP)^T M_P + M_P(\frac{1}{2}K)M_P = 0 \quad (4.64)$$

Now, notice that  $A + KP$  is guaranteed Hurwitz by our assumptions, and there is no state penalty matrix in (4.64). It is well known that such a Riccati equation has the unique stabilizing solution  $M_P = 0$ . Thus, what we have shown is that if  $P = P^T$  and  $M_P = M_P^T$ , then if  $P$  is a solution to the SDRE, Method 2 has a solution (for  $N = 0$ ) if and only if  $M_P = 0$ , and the solutions from the two methods are identical. What this in turn implies is that  $P$  is a constant matrix, so that we see that no matter which method we use to compute  $V$  we get

$$V = x^T P x = 2x^T P \int_0^1 tx dt \quad (4.65)$$

In summary, we have the following results. Assume  $P$  is a solution to the SDRE (4.58), so that  $P = P^T$ . Then

- if  $M_P = M_P^T$ , Method 2 is solvable and  $V$  is given by (4.60). Method 1 is solvable (with  $N = 0$ ) iff  $M_P = 0$ , so that  $P$  is a constant matrix, and both methods give the same answer
- if  $M_P \neq M_P^T$ , then neither method is solvable for  $V$

What this basically means is that Method 1 offers no advantages over Method 2 except possibly in the case where we allow  $N < 0$ , so that Method 2 is the preferred solution method.

To close this section, we observe that the above analysis depends on  $P$  being symmetric in Method 2, but we have seen in Chapter 3 that both methods can yield valid local solutions if  $P$  is allowed nonsymmetric. Section 4.3 proves that  $P$  can never be nonsymmetric in Method 1. Allowing nonsymmetric  $P$  in Method 2, however, eliminates the possibility of using standard numerical Riccati equation solvers, and thus mandates the analytical solution of the Riccati equation and PDE involving  $M_P$ . Research performed by this author indicates that analytical solutions to SDREs will be difficult to obtain in most problems of interest, and thus we shall consider  $P = P^T$  and focus on numerical Riccati solutions in the remainder of this dissertation.

#### 4.6 Optimal Control Numerical Solution Algorithm

As mentioned in the previous section, in the course of this research several failed attempts at analytical solutions to simple multistate example problems ( $n = 2$ ) were made, and we thus decided to focus on numerical solutions. We therefore developed and implemented a numerical SDRE solution approach to the optimal nonlinear regulator problem based on the results of Section 4.4. In [13], the authors suggest two potential ways of solving the SDRE necessary condition for optimality (4.42). One is to assume the SDC factorization parameter vector  $\alpha$  is an explicit function of time, and the other is to assume  $\alpha$  is an explicit function of the state vector  $x$ . They then outline an algorithm to solve (4.42) based on  $\alpha = \alpha(t)$ , which involves:

##### Procedure 1

- i. choosing an  $\alpha(0)$

- ii. solving the SDRE for  $P(x, \alpha)$
- iii. differentiating the SDRE with respect to the  $x_i$  and  $\alpha_i$  to obtain Lyapunov equations for the  $P_{x_i}$  and  $P_{\alpha_i}$
- iv. solving those Lyapunov equations
- v. substituting back into the necessary condition for optimality and algebraically solving to obtain  $\dot{\alpha}$
- vi. integrating over one time step and returning to step ii above until  $\alpha(t)$  diverges
- vii. iteratively adjusting  $\alpha(0)$  until  $\dot{\alpha}(t) \rightarrow 0$  as  $t \rightarrow \infty$ .

This is quite a cumbersome procedure having no hope of real time implementation since in the last step one must somehow compute how slight changes in  $\alpha(0)$  affect things farther and farther out in time, and change  $\alpha(0)$  accordingly. We propose instead a solution procedure similar to the above but based on combining both assumptions; that is,  $\alpha = \alpha(x, t)$ . At each time step we let each  $\alpha_i$  be expressed as a polynomial (of a fixed, user-desired order) in  $x$  with unknown constant (over one time step) coefficients  $\alpha_{il}$ . The proposed method is then as follows:

#### Procedure 2

- i. stack the  $\alpha_{il}$  into a new unknown vector  $\theta$
- ii. choose  $\theta_0$  ( $\theta_0 = [0 \dots 0]^T$  may be a good choice)
- iii. solve the SDRE for  $P(x, \theta)$
- iv. follow step iii in Procedure 1, substituting  $\theta_i$  for  $\alpha_i$
- v. differentiate the expressions from step iv above with respect to  $x_i$  and  $\theta_i$  to obtain Lyapunov equations for  $P_{x_i \theta_i}$  and  $P_{\theta_i^2}$  (the partial of  $P$  with respect to  $\theta_i$  twice)
- vi. use the information from the previous step to form the gradient of our simplified necessary condition for optimality (4.56) with respect to the unknown parameter vector  $\theta$

- vii. use an iterative root finding procedure such as Newton's method to calculate the next value of  $\theta$  (note: we may not need the previous step if we choose another method that does not require gradients)
- viii. return to step iii and iterate until  $\theta$  converges
- ix. use the obtained values of  $P$  and  $\alpha$  to form the closed loop system
- x. integrate to the next time step and repeat, using the last value of  $\theta$  as the new initial guess in step ii

Note that by using this procedure we have converted our optimal control problem into a parameter optimization problem to be solved at each time step. Also note that this procedure actually allows  $\alpha$  to be a function of both time and the state vector, since it is possible a time-varying  $\theta$  vector may be obtained. Such a methodology has at least two advantages over Procedure 1. First, we have a good initial guess for  $\theta$ , since any value of  $\theta$  with only zeros or ones in the constant coefficient parts comprises a valid parametrization which, if held constant over the duration of the simulation, will from experience yield suboptimal performance close to optimal. Second, we do not need to integrate the dynamics into the future to adjust  $\theta$  iteratively. This represents a major reduction in computational burden and a great simplification in implementation.

This procedure was implemented in Matlab/Simulink for Example 2 of [14], the parameters of which are given below:

$$\begin{aligned} \dot{x}_1 &= x_1 - x_1^3 + x_2 + u_1 \\ \dot{x}_2 &= x_1 + x_1^2 x_2 - x_2 + u_2 \end{aligned} \tag{4.66}$$

with  $H = I_2$  and  $R = 2I_2$ , where  $I_2$  is the two dimensional identity matrix. This is a second-order example, so that  $\alpha$  is a scalar (function of  $x$ ). Using only a constant term in the polynomial expansion for  $\alpha$ , we were able to duplicate the optimal and suboptimal control and state trajectories

given in [14], starting from the initial condition  $x_0 = [1 \ 1]^T$ , where for the suboptimal case we have

$$A(x) = A_2(x) = \begin{bmatrix} 1 - x_1^2 & 1 \\ 1 + x_1 x_2 & -1 \end{bmatrix} \quad (4.67)$$

These histories are given in Figures 4.1, 4.2, and 4.3, where the suboptimal trajectories have been specifically labeled with an 'so' suffix, and the neighboring optimal trajectories have been left unlabeled. Interestingly enough we did not obtain the same nonconstant  $\alpha$  time history as Cloutier *et al.* This seems to suggest possible nonuniqueness of  $\alpha$ . Initial condition tradeoff studies were also performed. With  $\alpha_0$  fixed, the simulation was repeated for various initial  $x$  conditions, in effect mapping out a small, elliptical shaped region of convergence centered at the origin of the state space for this algorithm. With  $x_0$  confined to this region, convergence was obtained for various values of  $\alpha_0 \in [-1.5, 1.5]$ , and nonconvergence for  $\alpha_0$  outside this region (both plus and minus). In the convergent cases very few intermediate Newton iterations were required. Another interesting phenomenon observed was that, by choosing small magnitude initial conditions, we did indeed obtain a constant (with respect to time) profile of  $\alpha$ , although the values obtained differed depending on the initial conditions chosen. This appears to support the idea of strong  $x$  dependency in  $\alpha$  and using a Taylor series expansion in  $x$  to represent  $\alpha$ , since constant values of  $\alpha$  were observed for very small  $x$  values. Finally, we explored the use of this method with the  $\alpha$  updating turned off. For all values of  $\alpha$  chosen (ranging from -10,000 to 10,000), stable trajectories which displayed the same type of behavior in not satisfying the necessary condition for optimality (4.56) were obtained. Virtually all profiles were smooth as in Figure 4.4, somewhat large in the beginning and converging rapidly to zero within a couple of seconds (the initial condition responses decayed to zero within about four seconds in all cases).

Although this algorithm proved successful in obtaining the optimal control within a small radius of convergence from the origin, simple attempts to extend the convergence radius failed. As the suboptimal trajectories differed only slightly from the optimal ones, the payoff in seeking optimal controls at the cost of a high increased workload was judged to be insufficient to warrant further effort

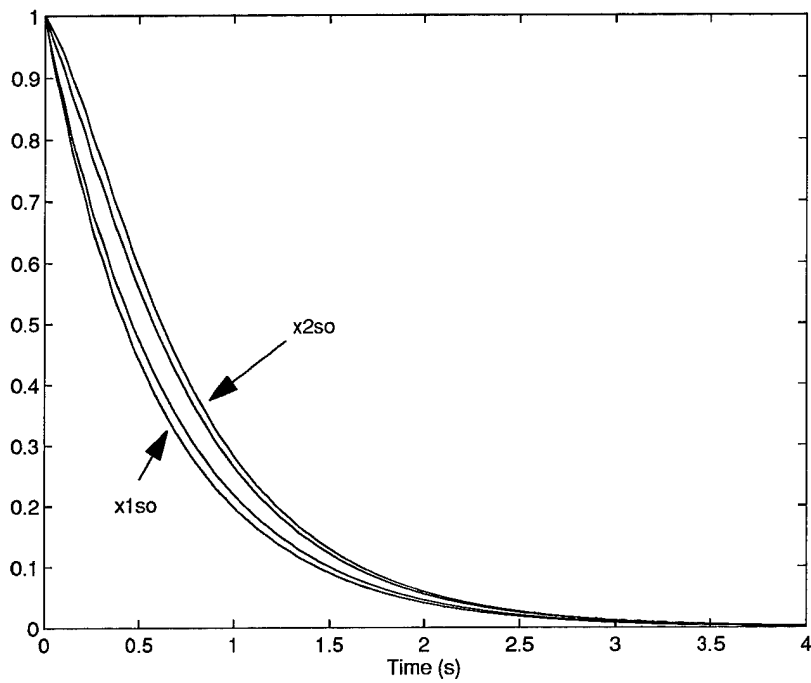


Figure 4.1: Optimal and Suboptimal State Histories

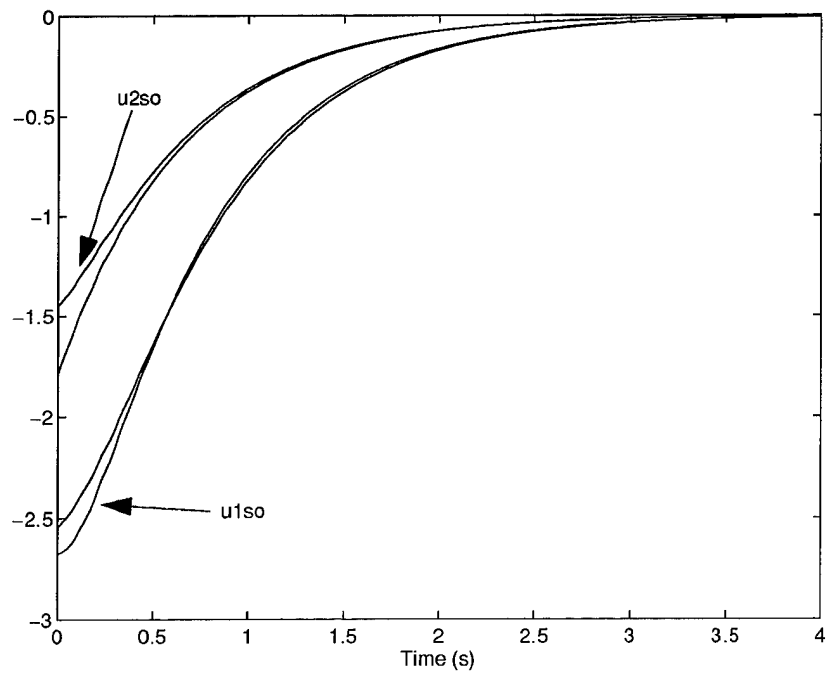


Figure 4.2: Optimal and Suboptimal Control Histories

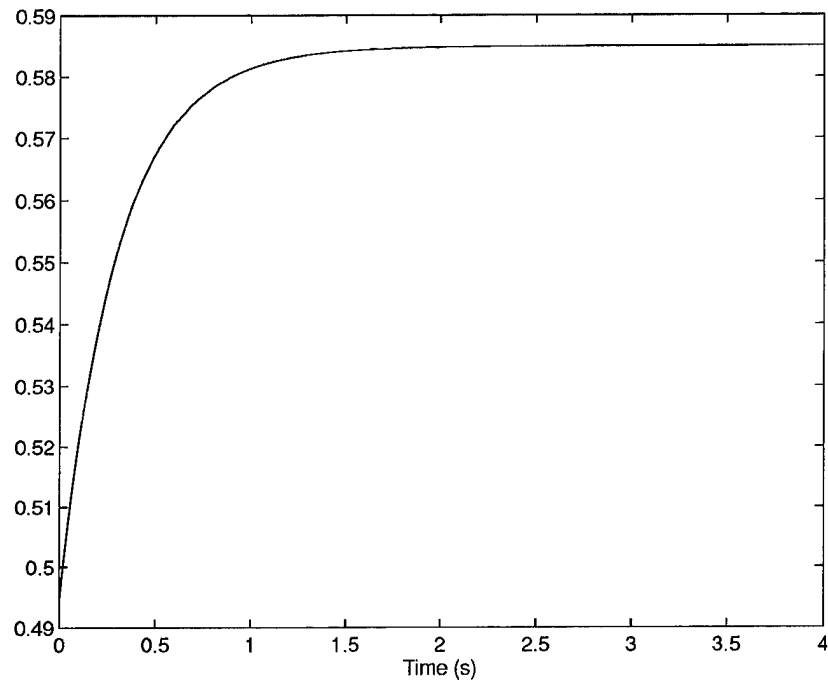


Figure 4.3: Optimal  $\alpha$  History

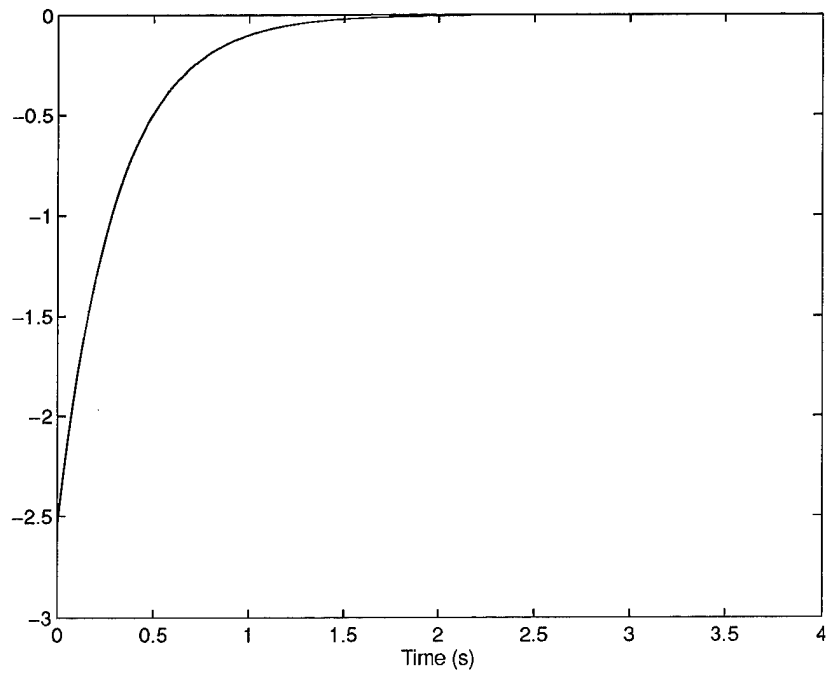


Figure 4.4: Necessary Condition for Optimality History

in this area. Suboptimal control strategies based on SDRE solutions for fixed SDC parametrizations were instead pursued.

#### 4.7 Solvability of *HJIs* and *HJB Equations*

In this section we give some brief qualitative comments regarding the solvability of *HJIs* and *HJB* equations in general, and in particular their solution by SDRE methods, where by solvable we mean that a satisfactory positive semidefinite storage (cost) function may be found which solves (2.125) or (2.135). We first give a summary of known results, and then proceed with some original observations.

Solution of *HJIs* is in general nontrivial and relatively few proposed solution techniques or solved problems can be found in the literature. In 1969, Lukes [47] presented an inductive power series approximate solution method for Hamilton-Jacobi-Bellman equations like (2.135). Van der Schaft extended this method to Hamilton-Jacobi-Issacs equations in [66]. A similar approximate power series solution method is proposed in [29]. In [71] and [72] an iterative power series approach was used to solve the state feedback Hamilton-Jacobi-Issacs equation approximately, where the quadratic term in  $V(x)$  and thus the linear part of the control was obtained via a gain-scheduled solution to the linearized problem. In [48], the method of characteristics was used to solve an  $H_\infty$  aerospace plane ascent problem numerically. Analytical nonlinear  $H_\infty$  solutions have been shown to exist for state feedback control of rigid spacecraft in [37] and [16], and an analytical output feedback solution was given for the very special case of passive systems with an assumed, restrictive structure in [15]. An output feedback solution for lossless systems of special structure was also given in [67]. In [45] a solution approach based on nonlinear matrix inequalities (NLMIs), which are actually state-dependent LMIs, is proposed. This approach is actually a generalization of the SDRE theory considered in this dissertation. This becomes clear when one examines [30], in which the NLMI approach is implemented via a finite difference scheme, and a single common, constant, local solution to multiple pointwise LMIs is sought. From this discussion it can be seen that the number of solved nonlinear  $H_\infty$  control problems is very small, and virtually every proposed method,

except the recent NLMI method, seeks solutions to the Hamilton-Jacobi-Issacs equation (HJIE), as opposed to the strict inequality form of the HJI. This observation motivates a brief discussion of the difference between solving strict HJIs

$$\mathcal{H}(V_x, x, u_*, d_*) < 0 \quad \forall x \neq 0 \quad (4.68)$$

versus solving the equality form (HJIE)

$$\mathcal{H}(V_x, x, u_*, d_*) = 0 \quad (4.69)$$

As mentioned in the above discussion, most solution attempts in the literature focus on (4.69). However, in [32], Imura *et al.* develop theory for the strict inequality case based on a strict bounded real condition for nonlinear systems. This is done to build internal stability directly into the problem, by requiring a positive definite HJI solution with strictly negative definite derivative (when  $d = 0$ ), which serves as a Lyapunov function to guarantee asymptotic stability of the closed loop system. The advantage gained from solving (4.68) (thus requiring negative definite  $\dot{V}$ ) is seen to be from Theorem 2.1.4 that all bounded closed loop solutions are guaranteed to converge to  $x = 0$ . Thus, boundedness implies closed loop asymptotic stability if (4.68) holds. Of course, when  $V > 0$  as in [32], Lyapunov's Theorem may directly be used instead of Lasalle's Invariance Principle to guarantee asymptotic stability. For the equality case, (4.69) just guarantees  $\dot{V} \leq 0$  when  $d = 0$ . Thus, this approach requires the use of Theorem 2.1.4 and its additional complexities to guarantee closed loop stability. It is easily seen from (2.126) that a negative definite  $\dot{V}$  can be recovered from (4.69) by simply requiring

$$h(x) = 0 \Leftrightarrow x = 0 \quad (4.70)$$

so that, when (4.70) holds, there is no advantage to solving the strict HJI (4.68). On the other hand, a strict HJI (4.68) can easily be transformed into the form of (4.69), if we modify the output  $h$  to be globally nonsingular as in (4.70). To illustrate, suppose we wish to solve (4.68). This may be accomplished by setting

$$\mathcal{H}(V_x, x, u_*, d_*) \leq -w(x) \quad (4.71)$$

where  $w(x)$  is a positive definite function of  $x$ . Thus,  $-w(x) < 0 \forall x \neq 0$ ,  $w(0) = 0$ , so that (4.71) and in turn (4.68) are satisfied if we solve the HJIE

$$\mathcal{H}(V_x, x, u_*, d_*) + w(x) = 0 \quad (4.72)$$

In terms of the SDRE method, if we have a fixed  $h$  and we desire to solve the strict inequality, we can instead solve the HJIE

$$A^T P + PA + PBKB^T P + \dot{Q} = 0 \quad (4.73)$$

where we have the new state weighting matrix function  $\dot{Q} = H^T H + W > 0$ , where  $w = x^T W x$  is any globally positive definite function, i.e.,  $W = \epsilon I$ ,  $0 < \epsilon \ll 1$ . Thus, solving (4.68) is essentially equivalent to solving an HJIE, where  $h^T h$  is restricted to be positive definite. In the SDRE context this requires  $Q = H^T H > 0 \forall x$ , which in turn implies  $P(x) > 0 \forall x$ . In solving (4.69), however, we can allow  $h^T h$  to be only positive semidefinite, thus opening up design options not available if we solve (4.68). Closed loop stability may be more complicated to guarantee, but the additional design flexibility may prove extremely useful. These issues are explored further in Chapters 8, 9, and 12.

We now move on to discuss issues common to solving HJIs, regardless of whether we choose to solve the strict inequality or the equality form. We consider both system type properties and algorithm type properties specific to the SDRE methods. In the first category we give a necessary condition for solvability of an HJI/HJB equation by any control strategy (which has to do with nonlinear controllability issues), and in the second category we are concerned with the effects of various choices of SDC parametrizations in the SDRE solution approaches. Letting  $d=0$ , we first consider the nonlinear regulator type problem. As per [33], a system

$$\dot{x} = a(x) + b(x)u$$

may locally be decomposed about any point of interest into state vector dynamics  $\dot{x}_1$  which are affected by the control, and state vector dynamics  $\dot{x}_2$  which are unaffected by the control, with dimensions depending on the dimension  $d_c$  of the smallest nonsingular involutive distribution  $\Delta_c$  invariant under  $a$  and the  $b_i$  and also containing the span of the  $b_i$  ( $\dim(x_1) = d_c$ ,  $\dim(x_2) = n - d_c$ ,

where  $n$  is the dimension of the state space). The decomposed system dynamics take the form

$$\begin{aligned}\dot{x}_1 &= a_1(x_1, x_2) + b_1(x_1, x_2)u \\ \dot{x}_2 &= a_2(x_2)\end{aligned}\tag{4.74}$$

Thus, if  $d_c < n$ , the  $x_2$  states will evolve according to  $a_2$  along some input-independent manifold in the state space. If the system does not converge to the origin along this manifold, then the overall system cannot be stabilized to the origin, regardless of what control is selected. Since finding a locally positive definite solution to the HJB equation for a zero state observable system implies a closed loop system which is stabilized to the origin, we obviously will not be able to find such a solution if such an unstable input-independent manifold exists. A necessary condition for local solvability of the nonlinear regulator SDRE is therefore that no such unstable input-independent manifold exists, or equivalently, that the system be **nonlinearly stabilizable**. The theory may be extended [33] to the global case by considering the span of the Control Lie Algebra instead of the distribution  $\Delta_c$  described above, an extension which also allows for singular distributions. Thus, prior to attempting to solve a given problem, we should consider the Control Lie Algebra for the problem, and determine if and where such uncontrollable, unstable manifolds exist. Note that this applies to nonlinear  $H_\infty$  as well as nonlinear regulation, since both require stabilization to the origin in the closed loop. These issues are explored in greater detail in Chapter 6. In addition to the system controllability properties, we require various assumptions on the SDC factorizations to guarantee solvability of HJI/HJB equations via the SDRE techniques. In the case of analytic systems describable by a single state ( $n = 1$ ), sufficient conditions for solvability are given in Chapter 5, as well as necessary conditions for most cases of interest. In the multistate case we are motivated by the results of Sections 4.5 and 4.6 to forego solving the HJI/HJB equations to obtain the optimal solution and associated cost function to use as a Lyapunov function, but instead to consider suboptimal solutions based on SDREs alone, seeking alternate forms of Lyapunov functions. The effects of SDC parameterizations on this process are explored in detail in Chapters 6, 7, 8, 9, and 10.

## V. Solution Properties for Scalar Analytic Systems

### 5.1 Introduction

In this chapter we examine the SDRE nonlinear regulator for the scalar (single-state) analytic case, giving necessary and sufficient conditions for obtaining (locally) stabilizing solutions for almost all possible sets of assumptions on dynamics, input matrices, and weighting functions. Although obviously not widely applicable, these results offer significant insight into the nature of SDRE solutions and their relationships with the linear theory.

As in Section 2.5.1, we consider regulation of input-affine nonlinear dynamical systems, but here we limit ourselves to systems describable by a single state variable,  $x$ . We assume perfect measurements and that the system has an equilibrium at the origin. For such a system we may write

$$\begin{aligned}\dot{x} &= a(x) + b(x)u, \quad a(0) = 0 \\ z &= \begin{bmatrix} h(x) \\ u \end{bmatrix}, \quad h(0) = 0\end{aligned}\tag{5.1}$$

where  $u$  is a scalar control,  $z$  is a scalar penalized variable, and  $a$ ,  $b$  and  $h$  are assumed analytic real-valued functions of  $x$ . The control objective is minimization of the cost function

$$J = \frac{1}{2} \int_0^\infty h^2 + u^2 dt\tag{5.2}$$

In [13] it is shown that the above objective may be accomplished by using the following nonlinear regulator SDRE technique:

- i. Write (5.1) in the so-called state-dependent coefficient (SDC) form

$$\begin{aligned}\dot{x} &= A(x)x + B(x)u \\ z &= \begin{bmatrix} H(x)x \\ u \end{bmatrix}\end{aligned}\tag{5.3}$$

ii. Solve the SDRE

$$A(x)p(x) + p(x)A(x) - p(x)B^2(x)p(x) + H^2(x) = 0 \quad (5.4)$$

iii. Construct the optimal state feedback via

$$u_o = -B(x)p(x)x \quad (5.5)$$

If the pairs  $\{A(x), B(x)\}$  and  $\{H(x), A(x)\}$  are respectively stabilizable and detectable for all  $x$ , where stabilizability/detectability are defined in the standard linear sense, then it has been shown [13] that the above algorithm gives a locally asymptotically stable closed loop system. In this chapter we assume analyticity of the system parameters  $a$ ,  $b$ , and  $h$  and seek additional conditions under which the SDRE control algorithm yields an *analytic* locally stabilizing state feedback. This line of inquiry is motivated by the fact that, if the matrices involved in an algebraic Riccati equation (ARE) are analytic in a parameter and stabilizability and detectability assumptions hold, then the maximal solution to the ARE has been shown also to be analytic in that parameter [58]. Thus, by assuming analyticity of the system parameters and of the stabilizing solution of (5.4), we expect to recover the stabilizability/detectability assumptions of [13], at least as a special case.

## 5.2 SDRE Solutions

For the single-state input-affine case considered here, the local solution to the state feedback non-linear regulator SDRE may be constructed explicitly. Dropping the  $x$  dependency notation and rearranging, the SDRE (5.4) becomes

$$B^2p^2 - 2Ap - H^2 = 0 \quad (5.6)$$

while using the optimal control (5.5), the closed loop system becomes

$$\dot{x} = Ax + B(-Bpx) = (A - B^2p)x \quad (5.7)$$

We now introduce some convenient notation. Recall that if  $g$  is a (scalar valued) analytic function of  $x$  on  $R$ , then  $g$  has a unique convergent Taylor series expansion about the origin

$$g(x) = g(0) + \frac{\partial g}{\partial x} \Big|_0 x + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \Big|_0 x^2 + \dots \quad (5.8)$$

so that near  $x = 0$ ,  $g$  acts as a polynomial

$$g(x) = g_0 + g_1 x + g_2 x^2 + \dots \quad (5.9)$$

where the coefficients  $g_i$  are simply the  $i$ th derivatives of  $g$  with respect to  $x$ , evaluated at  $x = 0$  (weighted by known, constant scalars). Using (5.1) we may thus write

$$a(x) = a_1 x + a_2 x^2 + \dots \quad (5.10)$$

$$h(x) = h_1 x + h_2 x^2 + \dots \quad (5.11)$$

so that  $A$  and  $H$  are given uniquely as

$$\begin{aligned} A(x) &= a_1 + a_2 x + a_3 x^2 + \dots \\ &= A_0 + A_1 x + A_2 x^2 + \dots \\ H(x) &= h_1 + h_2 x + h_3 x^2 + \dots \\ &= H_0 + H_1 x + H_2 x^2 + \dots \end{aligned} \quad (5.12)$$

and we see that  $A$  and  $H$  are themselves analytic functions of  $x$ . Now let  $d_g$  represent the minimum power of  $x$  in the polynomial expression for  $g$  having nonzero coefficient. Also, let that nonzero coefficient be  $c_g$ . As per the above notation, let  $d_A$ ,  $d_B$ ,  $d_H$  and  $d_p$  represent the minimum power of  $x$  with nonzero coefficient in the polynomial expressions for  $A$ ,  $B$ ,  $H$  and  $p$ , respectively. Also, let  $c_A$ ,  $c_B$ ,  $c_H$ , and  $c_p$  represent the associated nonzero coefficients. We now give a well known stability lemma, repeated from Section 2.1.3 for ease of reference [33].

**Lemma 5.2.1** *Consider the one-dimensional system*

$$\dot{x} = cx^m + Q_m(x)$$

with  $m \geq 1$ ,  $c \neq 0$ , and  $Q_m(x)$  a function vanishing at zero together with all partial derivatives of order less than or equal to  $m$ . The equilibrium  $x = 0$  is asymptotically stable if  $m$  is odd and  $c < 0$ . Otherwise it is unstable.

The general procedure in the following development is to make assumptions on the relationships of the various  $d_*$ , assume existence of an analytic solution  $p$ , group coefficients of the smallest contributing power of  $x$ , and invoke Lemma 5.2.1 to determine when stabilizing solutions exist. We shall consider four cases, which include several mutually exclusive subcases.

Case 1 ( $d_H < d_B$  and  $2d_H < d_A$ ) For this case the lowest order nonzero part of (5.6) is

$$-c_H^2 x^{2d_H} = 0 \quad (5.13)$$

which, since  $c_H$  is nonzero, has no solution for all  $x$ . Obviously, there is thus no stabilizing solution in this case. This result reinforces the intuitive feeling that we cannot penalize powers of  $x$  smaller than those on which we may have some effect (either through the control or through the dynamics themselves).

Case 2 ( $d_A < 2d_B$ ) With these assumptions and not allowing those of Case 1, solutions exist to (5.6) such that  $d_p = 0$  for  $2d_H = d_A$  and  $d_H = d_B$ . Otherwise  $d_p \geq 1$  and the existence of solutions is highly dependent on the particular structures of  $A$ ,  $B$ , and  $H$ . We thus give no criteria for the existence of stabilizing solutions in this case. Regardless, examining (5.7), we see that since  $d_p \geq 0$  in any case and that  $d_A < 2d_B$ , then the stability of the closed loop system is unaffected by the control, and is in fact determined by  $d_A$  and  $c_A$ . Invoking Lemma 5.2.1 we conclude for Case 2 that any solution that exists is stabilizing iff  $d_A$  is even and  $c_A < 0$ .

Case 3 ( $d_A = 2d_B$ )

Case 3A ( $d_H > d_B$ ) In this case,  $d_p = 0$  and (5.6) reduces to

$$c_p(c_B^2 c_p - 2c_A)x^{d_A} + \dots = 0 \quad (5.14)$$

Thus, the two possible solutions to (5.14) are  $c_p = 0$ ,  $c_p = 2c_A/c_B^2$ . For  $c_p = 0$ , the optimal control is  $u_o = 0$ , giving a closed loop system of  $\dot{x} = Ax$  (to 1st order). Thus, this solution is stabilizing iff the open loop system is stable ( $d_A$  is even and  $c_A < 0$ ). For  $c_p = 2c_A/c_B^2$ , the closed loop system becomes (to leading order)

$$\dot{x} = c_A x^{d_A+1} + c_B \left( -c_B \frac{2c_A}{c_B^2} x^{d_A+1} \right) = -c_A x^{d_A+1} \quad (5.15)$$

Thus, we see that (5.15) will be stable iff  $d_A$  is even and  $c_A > 0$ . Combining these two results we conclude that stabilizing solutions exist for this subcase iff  $d_A$  is even, which is always the case since  $d_A = 2d_B$ , and  $d_B$  is a nonnegative integer.

This subcase extends to the situation in which our cost function to be minimized is identically zero ( $d_H \rightarrow \infty$ ). We observe that the nonlinear behavior is analogous to that observed for linear time-invariant (LTI) systems in LQR theory. For LTI systems, if the open loop system is stable, the LQR optimal control for zero cost function is zero, just as we see here. If an open loop LTI system is unstable and controllable, then the open loop unstable poles are moved to their stable mirror images in the left half complex plane (i.e., the real parts of the eigenvalues of the closed loop system are the negatives of their open loop counterparts). In this nonlinear open loop unstable single-state case, we see that the closed loop dynamics are the negative of the open loop dynamics ( $a_{cl} = -a_{ol}$ ).

**Case 3B** ( $d_H = d_B$ ) In this case  $d_p$  again equals zero and the leading-order (positive) SDRE solution is given by

$$c_p = \frac{c_A}{c_B^2} + \sqrt{\frac{c_A^2}{c_B^4} + \frac{c_H^2}{c_B^2}} \quad (5.16)$$

Substituting (5.16) into  $u_o = -bpx$  and then into (5.1) we obtain the (leading-order) closed loop dynamics

$$\begin{aligned} \dot{x} &= c_A x^{d_A+1} - c_B^2 \left( \frac{c_A}{c_B^2} + \sqrt{\frac{c_A^2}{c_B^4} + \frac{c_H^2}{c_B^2}} \right) x^{d_A+1} \\ &= -\sqrt{c_A^2 + (c_B c_H)^2} x^{d_A+1} \end{aligned} \quad (5.17)$$

We only consider the positive solution because it is trivial to show that the negative square root solution always yields an unstable closed loop system. We thus see in this case that we get stabilizing solutions iff  $d_A$  is again even.

Case 4 ( $d_A > 2d_B$ )

Case 4A ( $d_H = d_B$ ) In this case  $d_p = 0$  and (5.6) becomes

$$(c_H^2 - c_B^2 c_p^2) x^{2d_B} = 0 \quad (5.18)$$

giving

$$c_p = |c_H/c_B| \quad (5.19)$$

so that the closed loop dynamics become

$$\dot{x} = -|c_B c_H| x^{2d_B+1} + \dots \quad (5.20)$$

Thus, a stabilizing solution exists iff  $2d_B$  is an even integer, which is always the case.

Case 4B ( $d_H > d_B$ ) In this case  $d_p \geq 1$  and the existence of solutions is highly dependent on the particular structures of  $A$ ,  $B$ , and  $H$ . We thus give no criterion for the existence of stabilizing solutions in this case. We also note that such a case is somewhat unusual, in that we would be penalizing only large powers of  $x$  without penalizing the smaller powers.

### 5.3 Discussion

We now investigate the assumptions required for local stability in [13] in the framework of the above theory. Recall that in [13] the stability proofs involved assuming linear stabilizability (controllability) of the pair  $\{A(x), B(x)\} \forall x$  and linear detectability (observability) of the pair  $\{H(x), A(x)\} \forall x$  in some neighborhood of the origin. For purposes of clarity we break this down into controllable/observable and stabilizable/detectable cases.

For the single-state controllable/observable case the above assumption translates to requiring  $\text{rank}(B(x)) = 1$  and  $\text{rank}(H(x)) = 1$  for all  $x$  in a neighborhood of the origin, or, in effect,  $B$  and

$H$  both having nonzero constant parts and no roots “close” to the origin. The controllable case can easily be seen always to yield stabilizing solutions according to the above theory as follows. For this case we have  $d_B = d_H = 0$  so that only cases 3B and 4A above are possible, depending on  $d_A$ . Regardless of the value of  $d_A$ , it is trivial to see that either case yields stabilizing solutions.

For the stabilizable/detectable case, applying the definition of linear stabilizability/detectability, we must have  $A < 0$  whenever  $B$  or  $H$  are equal to zero ( $A$  must have negative eigenvalues when not controllable or observable). Such an assumption allows  $d_H$  and  $d_B$  to be greater than zero, and also allows roots ‘close’ to the origin, as long as the value of  $A$  at those locations is negative. This interpretation of stabilizable/detectable always results in stable closed loop systems, as follows. If either  $B$  or  $H$  (or both) loses rank at  $x = 0$ , then  $A$  must have a negative constant part in order to meet the above definition of stabilizable/detectable. Thus,  $d_A = 0$  and  $c_A < 0$ . Only cases 2 and 3 are possible, either of which yields stabilizing solutions. The only other possible choice is for neither  $B$  nor  $H$  to lose rank at  $x = 0$ , but for both to lose rank only at nonzero  $x$  values or not at all. In this case  $d_B = d_H = 0$  which is the same as the controllable/observable case discussed above.

The results of this chapter thus verify the known result of analytic and stabilizable/detectable systems yielding analytic stabilizing solutions to (5.4). The results of this chapter also show, however, that stabilizability/detectability for all  $x$  are not necessary for existence of analytic stabilizing solutions in the scalar case, and illustrate under what conditions such solutions may or may not be obtained.

## 5.4 Examples

We now present three examples to illustrate the above theory, and to demonstrate the gap between the sufficient conditions for stability in [13] and the necessary and sufficient conditions for analytic stabilizing solutions derived herein.

### Example 1

$$\dot{x} = x^3 + xu; \quad h = cx^2 \tag{5.21}$$

This example is not controllable/observable, nor stabilizable/detectable in a neighborhood of the origin, so that the sufficient condition of [13] tells us nothing. Here we have  $A = x^2$ ,  $B = x$ , so that  $d_A = 2$ ,  $c_A = 1$ ,  $d_B = 1$ ,  $c_B = 1$ ,  $d_H = 1$ , and  $c_H = c$  and Case 3B applies. Since  $d_A = 2$  is even, we expect a stable closed loop solution. Solving (5.16) we find

$$c_p = 1 + \sqrt{1 + c^2} \quad (5.22)$$

The resulting low-order control is

$$u = -(1 + \sqrt{1 + c^2})x^3 \quad (5.23)$$

giving the closed loop system

$$\dot{x} = -\sqrt{1 + c^2}x^3 + \dots \quad (5.24)$$

which is indeed (locally) stable.

#### Example 2

$$\dot{x} = x^2 + xu; \quad h = cx^2 \quad (5.25)$$

Notice that all we have changed is  $a(x)$  from  $x^3$  to  $x^2$  so that again the sufficient condition of [13] tells us nothing. Here we have  $A = x$ ,  $B = x$ , so that  $d_A = 1$ ,  $c_A = 1$ ,  $d_B = 1$ ,  $c_B = 1$ ,  $d_H = 1$ , and  $c_H = c$  so that Case 2 applies. Since  $d_A = 1$  is odd, we expect an unstable closed loop solution. Solving (5.6) we find  $d_p = 1$  and  $c_p = c^2/2$ , so that  $p$  is not even locally positive semidefinite. The resulting low-order control is

$$u = -x\left(\frac{c^2}{2}x\right) = -\frac{c^2}{2}x^3 \quad (5.26)$$

giving the closed loop system

$$\dot{x} = x^2 - \frac{c^2}{2}x^4 + \dots \quad (5.27)$$

which as expected has stability properties unaffected by the control, and is clearly unstable.

#### Example 3

$$\dot{x} = x^2 + (x - 2)u; \quad h = x + x^2 \quad (5.28)$$

In this example we have altered Example 2 to have a controllable/observable *linearization*, which is known [67] to guarantee a locally stabilizing solution. We now have  $A = x$ ,  $B = x - 2$ , so that  $d_A = 1$ ,  $c_A = 1$ ,  $d_B = 0$ ,  $c_B = -2$ ,  $d_H = 0$ , and  $c_H = 1$  so that Case 4A applies, and we always get a stabilizing solution. From (5.20) and (5.28) we find the closed loop system

$$\dot{x} = -2x + x^2 + \cdots \quad (5.29)$$

which is indeed (locally) stable.

From these examples we see that the gap between the stabilizability/detectability conditions and the conditions derived herein mainly focuses on sufficiency for systems that have  $d_A > 0$  ( $A$  has a zero constant term), and necessity in all other cases since the stabilizability/detectability conditions do not address necessity at all.

## 5.5 Conclusion

We have derived necessary and sufficient conditions for existence of analytic stabilizing solutions to the nonlinear SDRE regulator problem in the input-affine single-state case. These conditions depend on the smallest contributing powers of  $x$  in the polynomial expressions for the analytic system parameters. Although the single-state nature of the results limits their applicability, they nevertheless offer significant insight into the nature of SDRE solutions in general and their relationships with LQR theory. In the succeeding chapters we return to the multistate, multivariable case and derive conditions for stability which extend the results of this chapter.

## VI. Controllability Issues in SDRE Control

### 6.1 Introduction

In this chapter we return to the general multistate case, and begin to fill some of the holes currently existing in the theory of the SDRE-based methods, by examining controllability issues in the context of the state feedback regulator problem. These issues have not been addressed in the literature, and have significant implications for global stability of SDRE-based nonlinear control algorithms. Recall that we are dealing with continuous time, state feedback, input-affine, autonomous nonlinear dynamic systems of the form

$$\begin{aligned} \dot{x} &= a(x) + b(x)u, \quad a(0) = 0 \\ z &= \begin{bmatrix} h(x) \\ \bar{R}(x)u \end{bmatrix}, \quad h(0) = 0 \end{aligned} \quad (6.1)$$

with state vector  $x \in \mathcal{R}^n$ , control vector  $u \in \mathcal{R}^m$ , penalized variable  $z \in \mathcal{R}^s$ , and nonsingular (for all  $x$ ) control penalty matrix function  $\bar{R}(x)$ . We assume  $a$  and  $h$  are globally at least  $C^1$  so that one can obtain a factored system representation of the form

$$\begin{aligned} \dot{x} &= A(x)x + B(x)u \\ z &= \begin{bmatrix} H(x)x \\ \bar{R}(x)u \end{bmatrix} \end{aligned} \quad (6.2)$$

Based on (6.2) recall that our suboptimal solution for the optimal regulation problem, i.e., driving the state to zero while simultaneously keeping the cost function

$$J = \int_0^\infty z^T z \, dt = \int_0^\infty x^T Q(x)x + u^T R(x)u \, dt \quad (6.3)$$

‘close’ to its optimal value, where we have defined  $R(x) = \bar{R}^T(x)\bar{R}(x) > 0$  and  $Q(x) = H^T(x)H(x) \geq 0 \, \forall x$ , is given by setting

$$u(x) = -R^{-1}(x)B^T(x)P(x)x \quad (6.4)$$

where  $P(x)$  is the maximal, stabilizing solution to the steady state continuous time state-dependent Riccati equation (SDRE)

$$A^T(x)P(x) + P(x)A(x) - P(x)B(x)R^{-1}(x)B^T(x)P(x) + Q(x) = 0 \quad (6.5)$$

Now, in order for the desired solution of (6.5) to exist for all  $x$ , we must assume, for example, that the pairs  $\{A(x), B(x)\}$  and  $\{H(x), A(x)\}$  are controllable and observable, respectively, for all  $x$ , where we employ the common definitions of controllability and observability from linear systems theory [76]. Of course, less restrictive assumptions such as stabilizability and detectability for all  $x$  would also be sufficient to ensure global existence of unique, stabilizing solutions to (6.5). However, since determining stabilizability requires determining controllability, in this chapter we seek first to understand the relationship between the ‘factored’ controllability assumed in order to guarantee existence of solutions of (6.5), and the true nonlinear controllability of the system (6.1). Note that although SDRE requires observability (detectability) of the factorization to guarantee existence of Riccati solutions, true nonlinear observability of the system is guaranteed by the full state feedback assumption. Thus, comparison between nonlinear and factored observability will not be included in this first examination of the state feedback case.

## 6.2 Factored versus True Controllability of Nonlinear Systems

Comparison of factored versus true nonlinear controllability is facilitated by considering controllability in terms of the dimension of invariant locally reachable and unreachable spaces. For controllable systems, the dimension of the locally reachable space must be  $n$ , the dimension of the state space. Thus, for linear time-invariant (LTI) systems

$$\dot{x} = Ax + Bu \quad (6.6)$$

the controllability of the system (6.6) is established [41] by verifying the well-known condition that the rank of the controllability matrix,  $M_{cl}$ , is equal to  $n$ , where

$$M_{cl} = [B \ AB \ A^2B \ \cdots \ A^{n-1}B] \quad (6.7)$$

For the SDC factored system (6.2), the above test for controllability generalizes to the rank test on the factored controllability matrix function  $M_{cf}(x)$

$$\text{rank}[M_{cf}(x)] = \text{rank}[B(x) \ A(x)B(x) \ A^2(x)B(x) \ \cdots \ A^{n-1}(x)B(x)] = n \ \forall \ x \quad (6.8)$$

while for the original input-affine nonlinear system, local controllability is characterized [33] at each  $x$  in terms of the dimension of the span of the smallest nonsingular and involutive distribution,  $\Delta_c(x)$ , containing the columns  $b_i$  of  $B(x)$ ,  $1 \leq i \leq m$ , and invariant under  $a$  and the  $b_i$ . This distribution assigns to each  $x \in \mathcal{R}^n$  a vector space, an open subset of which is reachable from the given point by using piecewise constant inputs. A sufficient condition [33] for the system (6.1) to be locally controllable at the point  $x$  is thus

$$\text{rank}[\Delta_c(x)] = n \quad (6.9)$$

and the system is said to be *weakly controllable* if (6.9) holds for all  $x$ . Note that if we desire to indicate that (6.9) holds for all  $x$  in a set  $S$ , then we shall say the system is weakly controllable on  $S$ . Now, we recall that invariance of the distribution  $\Delta_c$  with respect to  $a$  means that the Lie bracket of  $a$  with any vector field  $\tau \in \Delta_c$  is a vector field which is also in  $\Delta_c$ . This fact is used in [33] to suggest a recursive algorithm for generating  $\Delta_c$ , as follows

- i. Let  $\Delta_0 = \text{span}(B) = \text{span}(b_i)$
- ii. Let  $\Delta_1 = \Delta_0 + [a, b_i] + [b_j, b_i]$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq m$  where  $[a, g]$  indicates the Lie bracket of the vector fields  $a$  and  $g$  defined by

$$[a, g] = \frac{\partial g}{\partial x}a - \frac{\partial a}{\partial x}g \quad (6.10)$$

and the  $+$  indicates the subspace sum, i.e., the sum of the spans.

- iii. Let  $\Delta_k = \Delta_{k-1} + [a, d_j] + [b_i, d_j]$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$  where the  $d_j$  form a basis for  $\Delta_{k-1}$
- iv. Terminate when  $\Delta_{k+1} = \Delta_k$

It is shown in [33] that the above procedure gives  $\Delta_c = M_{cl}$  in the case of LTI systems. Also, it is proven that  $\Delta_{n-1}$  generated using the above algorithm always equals  $\Delta_c$  on an open and dense

subset of  $\mathcal{R}^n$ , and that if each of the  $\Delta_k$  above is nonsingular, then  $\Delta_{n-1} = \Delta_c$  for all  $x$ . Thus, it takes at most  $n - 1$  nontrivial steps of the above recursive procedure to give a picture of the system controllability, which, although not total, is correct on most of the set of interest.

In light of the above discussion we now make two important observations. First, if a system fails to pass the appropriate controllability rank test, then the uncontrollable subspace may be determined [33] by finding the annihilator of the appropriate matrix, i.e., the left nullspace of  $M_{cl}$ ,  $M_{cf}$  or  $\Delta_c$ . Second, the theory guarantees that the controllable and uncontrollable subspaces as determined from  $M_{cl}$  and  $\Delta_c$  are invariant subspaces for the appropriate systems, while, on the other hand, the controllable and uncontrollable subspaces determined from  $M_{cf}$  hold only for the single  $x$  value being considered. Thus, locally uncontrollable and invariant sets might or might not exist for the factored system, and further analysis would be warranted to determine such.

The above concepts give rise to controllable/uncontrollable system decompositions [41], [33]. For an LTI system with controllable subspace of dimension  $d$ , the system (6.6) may be partitioned

$$\begin{aligned}\dot{x}_1 &= A_{11}x_1 + A_{12}x_2 + B_1u \\ \dot{x}_2 &= A_{22}x_2\end{aligned}\tag{6.11}$$

where  $x_1 \in \mathcal{R}^d$  is in the controllable subspace  $\mathcal{C}_l$ ,  $x_2 \in \mathcal{R}^{n-d}$  is in the uncontrollable subspace  $\mathcal{U}_l$ , and  $\{A_{11}, B_1\}$  is a controllable pair. For the nonlinear system (6.1) with controllable subspace of dimension  $d$ , we similarly may write

$$\dot{x}_1 = a_1(x_1, x_2) + b_1(x_1, x_2)u\tag{6.12}$$

$$\dot{x}_2 = a_2(x_2)\tag{6.13}$$

where  $x_1 \in \mathcal{R}^d$  is in the controllable subspace  $\mathcal{C}_{nl}$ ,  $x_2 \in \mathcal{R}^{n-d}$  is in the uncontrollable subspace  $\mathcal{U}_{nl}$ , and (6.12) satisfies  $\text{rank}[\Delta_c] = d$ . Finally, (6.11) – (6.13) allow us to discuss the notion of stabilizability. Conceptually, a system is stabilizable if its uncontrollable part is stable. For LTI systems, stabilizability is equivalent to  $A_{22}$  in (6.11) being a Hurwitz matrix, while nonlinear stabilizability requires that the zero dynamics subsystem (6.13) be stable. For the SDC factored

system of SDRE control (6.2), we have a decomposition of the form (6.11) for each  $x \in \mathcal{R}^n$ . As discussed above, this gives a pointwise description of controllable and uncontrollable states. However, the fact that the decomposition changes in general from point to point without guaranteed invariance greatly complicates stability analysis based on the decomposed subsystems.

### 6.3 Theorems and Examples

With some minor extensions, the discussion of Section 6.2 allows us to prove some elementary theorems, the first of which is the lack of a general equivalency between factored and true nonlinear controllability.

**Theorem 6.3.1** *Consider the system (6.1) with  $a(x)$  and  $h(x)$  assumed to be  $C^1$  functions, so that (6.1) may be written as in (6.2). Assume the pair  $\{A(x), B(x)\}$  is controllable for all  $x$ , so that (6.8) holds. Then the system (6.1) is not necessarily weakly controllable.*

**Proof:** The proof is by counterexample. Consider the system

$$\begin{aligned}\dot{x}_1 &= x_1 x_2 + x_2 \\ \dot{x}_2 &= u\end{aligned}\tag{6.14}$$

Choose

$$A(x) = \begin{bmatrix} x_2 & 1 \\ 0 & 0 \end{bmatrix}\tag{6.15}$$

and observe that  $b(x) = [0 \ 1]^T$ . Then

$$M_{cf} = [b \ A(x)b] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\tag{6.16}$$

so that  $M_{cf}$  is full rank for all  $x$ , and the assumptions of the theorem hold. Let us now construct  $\Delta_c$  according to the recursive procedure given in Section 6.2. We have

$$\Delta_0 = \text{span}[b] = \text{span} \begin{bmatrix} 0 \\ 1 \end{bmatrix}\tag{6.17}$$

$$\Delta_1 = \Delta_0 + [a, b] + [b, b] \quad (6.18)$$

Now, since  $[b, b] = 0$  for any vector field, and  $b$  is a constant vector so that  $\frac{\partial b}{\partial x} = 0$ , we have

$$\Delta_1 = \Delta_0 + [a, b] = b + \frac{\partial a}{\partial x} b = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} x_2 & x_1 + 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (6.19)$$

so that

$$\Delta_1 = \text{span} \begin{bmatrix} 0 & x_1 + 1 \\ 1 & 0 \end{bmatrix} \quad (6.20)$$

which equals  $\Delta_c$  on at least an open and dense subset of  $\mathcal{R}^2$ . Performing one more iteration of the procedure we find (noting that  $d_1 = b$ )

$$\Delta_2 = \Delta_1 + [a, b] + [a, d_2] + [b, b] + [b, d_2] \quad (6.21)$$

where

$$d_2 = \begin{bmatrix} x_1 + 1 \\ 0 \end{bmatrix} \quad (6.22)$$

Now, since  $[b, b] = 0$  and  $[a, b]$  is already in  $\Delta_1$ , we need only compute the other two Lie brackets in (6.21). It is easily established that

$$\begin{aligned} [a, d_2] &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 x_2 + x_2 \\ 0 \end{bmatrix} - \begin{bmatrix} x_2 & x_1 + 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 + 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} x_1 x_2 + x_2 - x_1 x_2 - x_2 \\ 0 - 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned} \quad (6.23)$$

and

$$[b, d_2] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (6.24)$$

so that we have  $\Delta_2 = \Delta_1 = \Delta_c$  and the procedure terminates. Thus, clearly, at  $x_1 = -1$ , we have  $\text{rank}[\Delta_c] = 1 \neq n = 2$ , so that (6.9) fails to hold and the theorem is proved. ■

From the proof of the previous theorem we see that it is possible for a state dependent factorization to hide the existence of an uncontrollable, invariant set (the set of all  $x \in \mathcal{R}^2$  such that  $x_1 = -1$

for (6.14)). This fact is not mentioned in [13], nor, it appears, has it historically been well known when SDRE type methods were previously suggested. In fact, in [70], a state-dependent regulation algorithm is given, and the factorization (6.15) is recommended as a better factorization for (6.14) than the choice

$$A(x) = \begin{bmatrix} 0 & x_1 + 1 \\ 0 & 0 \end{bmatrix} \quad (6.25)$$

because (6.15) allows solution for the control at all  $x$ , while (6.25) does not. More recently, [21] suggests an SDRE-like regulation algorithm accompanied by the assertion that it works well quite often, but with virtually no theoretical basis to support the claim, or to indicate when the method might fail.

Upon comparing  $M_{cf}$  and  $\Delta_c$  for (6.14), it can be seen that the two differ only in their second columns, because the second column of  $M_{cf}$  is given by  $A(x)b$ , while the second column of  $\Delta_c$  is  $J(x)b$ , where  $J$  is the Jacobian of  $a$ , i.e.,  $J(x) = \frac{\partial a}{\partial x}$ . This observation allows us to prove an equivalency relationship between factored and true nonlinear controllability for the special case of 2nd-order systems with constant  $B$  matrices, as formalized below.

**Theorem 6.3.2** *Consider the system (6.1) written as (6.2) with  $n = 2$ , and let  $B$  be a constant matrix. Also, assume that  $A(x)$  is chosen such that  $J(x)B = kA(x)B \forall x$ , where  $0 \neq k \in \mathcal{R}$ . Then, if the factorization (6.2) is controllable for all  $x$ , the system (6.1) is weakly controllable on an open and dense subset of  $\mathcal{R}^2$ . Conversely, if (6.1) is weakly controllable on  $\mathcal{R}^2$ , then (6.2) is controllable for all  $x$ .*

**Proof:** The proofs follow by simply computing and checking the rank of  $M_{cf}$  and  $\Delta_c$  under the given assumptions. For the first claim, with (6.2) being controllable for all  $x$ , (6.8) implies

$$\text{rank}[M_{cf}(x)] = \text{rank}[B \ A(x)B] = 2 \forall x \quad (6.26)$$

Now, for (6.1), on an open and dense subset  $V$  of  $\mathcal{R}^2$  we have

$$\Delta_c(x) = \Delta_1(x) = B + [a, b_i] + [b_j, b_i], \quad 1 \leq i \leq m, 1 \leq j \leq m \quad (6.27)$$

Now, since the  $b_i$  are constant vectors, we have  $[b_j, b_i] = 0 \forall i, j$  and  $[a, b_i] = -Jb_i$ , so that

$$\Delta_c(x) = [B \ J(x)B] = [B \ kA(x)B] \quad (6.28)$$

which is guaranteed to be full rank for all  $x$  since  $M_{cf}(x)$  is full rank and  $k \neq 0$ . The second claim follows by the reverse logic path: starting with  $\Delta_c(x) = [B \ J(x)B]$  assumed full rank for all  $x$ , and substituting for  $J(x)B$  to obtain (6.28), this implies  $M_{cf}(x)$  must be full rank for all  $x$ , since the nonzero scaling  $k$  may be pulled out of the second column of  $M_{cf}$ . ■

We note that the forward claim could be strengthened to weak controllability on all of  $\mathcal{R}^2$  by requiring the additional assumption that  $\Delta_c(x) = \Delta_1(x) \forall x$ , and not just on an open and dense subset of  $\mathcal{R}^2$ . In Theorem 6.3.2 we restrict attention to 2nd-order systems, because for higher-order systems we will, in general, need to perform more iterations on  $\Delta_k$  (as many as  $n - 1$  total before we know  $\Delta_c$  on an open and dense subset), which will require repeated Lie brackets. As shown below, the first Lie bracket computation yields the second entry of  $\Delta_c$ , which differs from the second entry of  $M_{cf}$  by only a single additive term, due to the fact that  $B$  is constant. The divergence between succeeding entries in the two matrix functions will increase as more Lie brackets are required, and would be significantly greater even at the first Lie bracket computation if  $B$  were not constant. We now illustrate these concepts by computing the needed Lie brackets, observing that since  $a(x) = A(x)x$ , then

$$J(x) = A(x) + \frac{\partial A}{\partial x}x = A(x) + T_A(x)x \quad (6.29)$$

where the partial derivative term  $T_A$  is a third rank tensor, and thus

$$[a, b_i] = Jb_i = Ab_i + T_Axb_i \quad (6.30)$$

Note first that if  $b_i$  were not constant, then the Lie bracket  $[a, b_i]$  would include another term from the partial of  $b_i$  with respect to  $x$ . Also note that from (6.30), for the conditions of the theorem to hold, we need  $T_A(x)xB = (J - A)B = (k - 1)AB$  where  $k$  cannot equal 0, and thus we cannot have  $T_A(x)xB = -AB$  for any  $x$  if we require the second entry in the controllability matrices to

provide some of the needed rank. We also see that the theorem requires that  $AB = 0$  if and only if  $JB = 0$ , and this is also a necessary condition if  $B$  is not full rank. Now, if, for example,  $n = 3$ , then a second Lie bracket computation needed to compute  $\Delta_2$  gives

$$[a, Jb_i] = \frac{\partial Jb_i}{\partial x}a - J^2b_i \quad (6.31)$$

and we see that although the assumptions of the theorem give  $A^2B$  in the range of  $J^2B$ , there is still potential mismatch between  $M_{cf}$  and  $\Delta_c$  due to the contribution from the term involving the partial of  $Jb_i$  with respect to  $x$ . Finally, we note that the conditions of the theorem are sufficient, but far from necessary. It is possible for both  $M_{cf}$  and  $\Delta_c$  to be full rank without  $AB$  being in the range of  $JB$  alone as the theorem requires, as long as both  $AB$  and  $JB$  provide the remainder of a spanning set of  $\mathcal{R}^2$  that is not provided by  $B$ .

The above expression (6.29) relating the Jacobian  $J$  of  $a$  in the original nonlinear system (6.1) to the factorization  $A$  in (6.2) allows us to draw an additional conclusion regarding factored and local nonlinear controllability, as stated in the following theorem.

**Theorem 6.3.3** *Consider the system (6.1) written as (6.2), and assume (6.8) holds. Then the system (6.1) is weakly controllable on some local neighborhood of the origin.*

**Proof:** If (6.8) holds for all  $x$ , then clearly (6.8) holds for  $x = 0$ . Now, observe that (6.29) implies that  $J(0) = A(0)$ , so that (6.8) gives

$$\text{rank}[M_{cf}(0)] = \text{rank}[B(0) \ J(0)B(0) \ J^2(0)B(0) \ \dots \ J^{n-1}(0)B(0)] = n \quad (6.32)$$

Now, it is well known [39, 62, 68] that in some local neighborhood of the origin, the behavior of the nonlinear system (6.1) is governed by that of its *linearization* about the origin (provided the linearization is nontrivial), which is an LTI system given by

$$\dot{x} = J(0)x + B(0)u \quad (6.33)$$

Clearly, (6.32) implies that (6.33) is controllable, so that the nonlinear system (6.1) is weakly controllable on the set where the linearization (6.33) dominates, and the theorem is proven. ■

The previous three theorems dealt with an equivalence relationship between  $M_{cf}$  and  $\Delta_c$ , when the computation of  $\Delta_c$  required computation beyond  $\Delta_0$ . If the full rank requirement is met by the  $B$  matrix itself, then we clearly have equivalence between factored and nonlinear controllability without any assumptions on  $A(x)$  or on the dimension of the state, as formalized in the following theorem.

**Theorem 6.3.4** *Consider the system (6.1) written as (6.2) , and assume  $B(x)$  has rank  $n$  for all  $x$ . Then (6.2) is controllable for all  $x$ , and the system (6.1) is weakly controllable on  $\mathcal{R}^n$ .*

**Proof:** The proof follows trivially from the construction of  $M_{cf}$  and  $\Delta_c$ , noting that  $B(x)$  appears as the first component of both matrix functions, and itself satisfies the rank requirement. ■

It is important to note that, for the assumptions of Theorem 6.3.4 to hold, we must have  $m \geq n$  (at least as many controls as states). We also observe that Theorem 6.3.4 includes the special case of  $B$  equal to a constant, full rank matrix. This fact is exploited in Chapter 7. We now illustrate the above theorems by means of some examples.

### Example 1

Consider the system

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1 x_2 + u\end{aligned}\tag{6.34}$$

We have

$$a = \begin{bmatrix} x_2 \\ x_1 x_2 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\tag{6.35}$$

and  $b$  is not full rank for any  $x$ , so that Theorem 6.3.4 may not be used. However, we have

$$J = \begin{bmatrix} 0 & 1 \\ x_2 & x_1 \end{bmatrix}\tag{6.36}$$

and

$$\Delta_c = \Delta_I = \begin{bmatrix} 0 & 1 \\ 1 & x_1 \end{bmatrix}\tag{6.37}$$

which has rank 2 for all  $x$ . Now, choose

$$A(x) = \begin{bmatrix} 0 & 1 \\ 0 & x_1 \end{bmatrix} \quad (6.38)$$

Then

$$A(x)B = \begin{bmatrix} 1 \\ x_1 \end{bmatrix} = JB \quad \forall x \quad (6.39)$$

so that the conditions of Theorem 6.3.2 are satisfied. Checking the factored controllability matrix we find

$$M_{cf}(x) = \begin{bmatrix} 0 & 1 \\ 1 & x_1 \end{bmatrix} = \Delta_c(x) \quad \forall x \quad (6.40)$$

and we see that the original system is weakly controllable on  $\mathcal{R}^2$ , while the factored system is controllable for all  $x \in \mathcal{R}^2$  as well. On the other hand, choose

$$A(x) = \begin{bmatrix} x_2 & 1 - x_1 \\ 0 & x_1 \end{bmatrix} \quad (6.41)$$

which is easily verified to yield  $A(x)x = a(x)$ . However,

$$M_{cf}(x) = \begin{bmatrix} 0 & 1 - x_1 \\ 1 & x_1 \end{bmatrix} \quad (6.42)$$

which clearly loses rank at  $x_1 = 1$ . Given the analysis above, the first choice of  $A(x)$  for this example is preferable, because it guarantees global existence of the control. From this example we thus see that even when the original system is weakly controllable, care must be taken when choosing the factorization, as both poor and good choices from an implementation standpoint may exist.

In the next example we illustrate the nonnecessity of the  $AB = kJB$  condition as discussed above.

### **Example 2**

Consider the system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1^2 + x_2^2 + u \end{aligned} \quad (6.43)$$

We have

$$a = \begin{bmatrix} x_2 \\ x_1^2 + x_2^2 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (6.44)$$

so that again Theorem 6.3.4 may not be invoked. Computing the Jacobian of  $a$  we find

$$J = \begin{bmatrix} 0 & 1 \\ 2x_1 & 2x_2 \end{bmatrix} \quad (6.45)$$

so that

$$\Delta_c = \Delta_1 = \begin{bmatrix} 0 & 1 \\ 1 & 2x_2 \end{bmatrix} \quad (6.46)$$

which has rank 2 for all  $x$ . Now, choose

$$A(x) = \begin{bmatrix} 0 & 1 \\ x_1 & x_2 \end{bmatrix} \quad (6.47)$$

Then

$$A(x)B = \begin{bmatrix} 1 \\ x_2 \end{bmatrix} \quad (6.48)$$

which does not equal a multiple of  $JB$  for any  $x$ . However, checking the factored controllability matrix we find

$$M_{cf}(x) = \begin{bmatrix} 0 & 1 \\ 1 & x_2 \end{bmatrix} \quad (6.49)$$

which is full rank for all  $x$ , and again we see that the original system is weakly controllable on  $\mathcal{R}^2$ , while the factored system is controllable for all  $x \in \mathcal{R}^2$ . Thus, even though the conditions of Theorem 6.3.2 are not satisfied,  $[1 \ 2x_2]^T$  and  $[1 \ x_2]^T$  both provide vector functions which, together with  $b = [0 \ 1]^T$ , span  $\mathcal{R}^2$  for any value of  $x$ .

Finally, we return to the counterexample in the proof of Theorem 6.3.1, and see how the discussion of necessary conditions for equivalence of factored and true nonlinear controllability in 2nd-order systems with constant  $B$  matrices following Theorem 6.3.2 comes into play.

### Example 3

Consider the system

$$\begin{aligned}\dot{x}_1 &= x_1 x_2 + x_2 \\ \dot{x}_2 &= u\end{aligned}\tag{6.50}$$

We have

$$a = \begin{bmatrix} x_1 x_2 + x_2 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\tag{6.51}$$

so that  $\text{rank}(B) = 1 \forall x$  and again Theorem 6.3.4 may not be invoked. Computing the Jacobian of  $a$  we find

$$J = \begin{bmatrix} x_2 & x_1 + 1 \\ 0 & 0 \end{bmatrix}\tag{6.52}$$

so that

$$\Delta_c = \begin{bmatrix} 0 & x_1 + 1 \\ 1 & 0 \end{bmatrix}\tag{6.53}$$

which has rank 1 at  $x_1 = -1$  as we saw before. Now, also as before choose

$$A(x) = \begin{bmatrix} x_2 & 1 \\ 0 & 0 \end{bmatrix}\tag{6.54}$$

Then

$$A(x)B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\tag{6.55}$$

which gave the factored controllability matrix

$$M_{cf}(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\tag{6.56}$$

which is full rank for all  $x$ . However, we see that  $kJB$  does not equal  $AB$  for  $x = -1$  for any  $k \neq 0$ . Also,  $T_A x B = (J - A)B = -AB$  when  $x_1 = -1$ , and we have  $AB \neq 0$  when  $JB = 0$ , both of which we noted were necessary conditions for equivalency of the controllability tests if  $B$  was not full rank. Thus, we observe the discrepancy between the factored system being controllable for all

$x \in \mathcal{R}^2$ , while the original system fails to be nonlinearly controllable on  $x_1 = -1$ . This highlights the reverse of what was found at the end of Example 1. There, we saw existence of nonglobally controllable factorizations for a weakly controllability system, while here we observe existence of globally controllable factorizations when the original system is not weakly controllable. Finally, this example illustrates the potential weakness of Theorem 6.3.3. Clearly, the controllability equivalence based on linearization does not hold in this example beyond a ball of radius one centered at the origin. By replacing the first element of  $a$  in (6.51) with  $x_1x_2 + kx_2$ , with  $k \in \mathcal{R}$ , and decreasing the absolute value of  $k$ , we can construct an example for which the conclusion of Theorem 6.3.3 holds on an arbitrarily small neighborhood of the origin. Thus, the study of global as opposed to just local controllability equivalence is indeed well-motivated.

#### 6.4 *Summary and Conclusions*

We have shown that, in general, nonlinear controllability of input-affine systems (6.1) and controllability of a state-dependent factorization (6.2) of such systems in terms of linear type tests are not equivalent concepts. We have shown that controllability of a state-dependent factorization for all  $x$  is sufficient to guarantee weak controllability on the neighborhood of the origin dominated by the linearized dynamics, and we have shown that if  $B(x)$  is full rank for all  $x$ , then the conditions for both controllability tests are simultaneously met, regardless of the choice of  $A$  matrix factorization or dimension of the state vector. We have also shown sufficient conditions for equivalency in the special case of second-order systems with constant  $B$  matrices, which may be extended to necessary conditions when  $B$  is not full rank. For higher order systems, or for systems with nonconstant, non-full rank  $B$  matrix functions, we have shown that conditions guaranteeing equivalency become increasingly complex to meet, due to differences between the Lie brackets which characterize nonlinear controllability and the  $A^jB$  products in the factored controllability test.

While factored controllability for all  $x$  is sufficient to guarantee global well-posedness of SDRE-based control algorithms, it is in general not a sufficient assumption to guarantee true controllability

of the nonlinear system outside some possibly small neighborhood of the origin, a fact which has significant ramifications for global stability of SDRE control. In fact, (6.12) and (6.13) indicate that, regardless of how well the SDRE algorithm operates, it can only affect the portion of the state which is nonlinearly controllable. Thus a necessary condition for stability of the SDRE regulator is that the nonlinear system be nonlinearly stabilizable, as discussed and defined in Sections 4.7 and 6.2. This may be sufficient if only regulation is to be attempted, but full nonlinear controllability may be required for trajectory tracking or model following type tasks.

Finally, in this chapter we have been concerned only with the controllability part of the controllability/observability conditions assumed for well-posedness of the SDRE algorithm, and our study of factored and true nonlinear controllability was motivated by the desire to analyze and predict possible success of the state feedback SDRE regulation algorithm in guaranteeing global asymptotic stability of the closed loop. Thus, it is fitting that we conclude by making some brief comments on the observability part of the well-posedness assumptions. If we limit ourselves to the state feedback case, we have assumed full nonlinear observability, so there would be no worry about factored observability hiding truly unobservable modes. Factored observability will, however, still play a part in stability analysis of the regulator, since pointwise unobservable modes may not receive any control action [41]. These issues are fully addressed in Chapter 9. In the output feedback case, additional theory of factored versus nonlinear observability, analogous to what has been derived herein, will be required for a full understanding of the SDRE-based methods. As mentioned previously, these issues are not addressed in this dissertation, but are proposed as avenues of further research.

## VII. Stability of Systems with Full Rank Constant $B$ Matrices

### 7.1 Introduction

In this chapter we build on the results of Chapter 6, in particular making use of the equivalence of factored and nonlinear controllability for systems with full rank, constant  $B$  matrices, as given in Theorem 6.3.4. We still consider continuous time, state feedback, input-affine, autonomous nonlinear dynamic systems described by 6.1 and 6.2, but we now assume we have state and control vectors  $x, u \in \mathcal{R}^n$ , penalized variable  $z \in \mathcal{R}^{2n}$ , and nonsingular constant control penalty matrix  $\tilde{R}$ . We assume that  $a$  and  $h$  are real-valued  $C^1$  functions of  $x$  on  $\mathcal{R}^n$ , so that  $A$  and  $H$  are chosen to be (at least) continuous functions. We further assume that  $b(x) = B$  is a full rank constant matrix and that  $H^T(x)H(x)$  is positive definite for all  $x$  ( $H(x)$  is globally nonsingular). Under these conditions we clearly have that the pairs  $\{A(x), B\}$  and  $\{H(x), A(x)\}$  are controllable and observable, respectively, for all  $x$ , where we employ the common definitions of controllability and observability from linear systems theory. Finally, we assume that  $x = 0$  is the only open loop equilibrium of (6.1). Using these assumptions we will now show that the continuous time state feedback nonlinear SDRE regulator yields a closed loop system which is globally asymptotically stable. We also show that the assumption of a single open loop equilibrium at the origin may be relaxed by placing some constraints on the magnitude and/or structure of the state weighting matrix function  $Q(x) \equiv H^T(x)H(x)$ .

### 7.2 Control Algorithm

Recall that the control is given by

$$u(x) = -R^{-1}B^T P(x)x \quad (7.1)$$

where  $P(x)$  is the positive definite maximal solution to the steady state continuous time state-dependent Riccati equation (SDRE)

$$A^T(x)P(x) + P(x)A(x) - P(x)BR^{-1}B^T P(x) + Q(x) = 0 \quad (7.2)$$

which is guaranteed to exist for all  $x$  due to the global controllability/observability of the SDC factorizations. Note that in (7.1) and (7.2)  $R$  is a constant matrix  $R = \bar{R}^T \bar{R} > 0$ , and  $B$  is also a constant matrix, but  $P$  remains a function of  $x$  due to the  $x$ -dependency of  $A$  and  $Q$ .

### 7.3 Global Asymptotic Stability via the Direct Method of Lyapunov

We now prove global asymptotic stability of the closed loop system resulting from the SDRE continuous time control algorithm outlined in Section 7.2. The method of proof involves standard Lyapunov analysis [36] in a transformed set of coordinates, and is formalized below.

**Theorem 7.3.1** *Consider the system (6.1) with  $a(x)$  and  $h(x)$  assumed to be  $C^1$  functions and such that  $a(0) = 0$ ,  $h(0) = 0$ . Assume further that  $a(x) = 0 \Rightarrow x = 0$ ,  $\text{rank}(B) = n$  and  $H^T(x)H(x) > 0 \forall x$  where  $h(x) = H(x)x$ . Also, assume  $\bar{R}$  in (6.1) is constant and nonsingular. Then application of the SDRE nonlinear regulator control algorithm defined by (7.1) and (7.2) to (6.1) yields a closed loop system which is globally asymptotically stable.*

**Proof:** As explained in Section 7.1, the assumptions on  $a(x)$  and  $h(x)$  guarantee that (6.1) may be written in the form of (6.2), where  $A(x)$  is nonsingular for all  $x$  since  $a(x) = A(x)x = 0$  only for  $x = 0$ . Consider now the change of coordinates defined by

$$q = \bar{B}^{-1}x \quad (7.3)$$

where

$$\bar{B} \equiv B\bar{R}^{-1} \quad (7.4)$$

and note that (7.3) and (7.4) define a globally valid static linear diffeomorphism since  $\bar{B}$  is a constant matrix,  $B$  has full rank, and  $\bar{R}$  is nonsingular. Consider now the Lyapunov function

$$V(q) = \frac{1}{2}q^T q \quad (7.5)$$

which is positive definite and decrescent, and also radially unbounded. Taking the derivative of (7.5) we find

$$\dot{V} = \frac{1}{2}\dot{q}^T q + \frac{1}{2}q^T \dot{q} \quad (7.6)$$

where from (7.3) we have

$$\dot{q} = \bar{B}^{-1}\dot{x} \quad (7.7)$$

Now, substituting (7.1) into (6.2) and using  $R^{-1} = (\bar{R}^T \bar{R})^{-1}$  we have

$$\dot{x} = A(x)x - BR^{-1}B^T P(x)x = A(x)x - B\bar{R}^{-1}\bar{R}^{-T}B^T P(x)x \quad (7.8)$$

so that using (7.4) we obtain

$$\dot{x} = A(x)x - \bar{B}\bar{B}^T P(x)x \quad (7.9)$$

Premultiplying (7.9) by  $\bar{B}^{-1}$  and substituting into (7.7) we find

$$\dot{q} = \bar{B}^{-1}A(x)x - \bar{B}^T P(x)x \quad (7.10)$$

Now, if we use the fact that  $x = \bar{B}q$  and define

$$A(q) \equiv \bar{B}^{-1}A(x)\bar{B} \quad (7.11)$$

and

$$P(q) \equiv \bar{B}^T P(x)\bar{B} \quad (7.12)$$

then (7.10) may be written

$$\dot{q} = A(q)q - P(q)q = [A(q) - P(q)]q \quad (7.13)$$

Now, observe that since (7.11) just defines a similarity transformation,  $A(q)$  has the same eigenvalues as  $A(x)$ , and thus  $A(q)$  is nonsingular for all  $q$ . Likewise [42], the congruence transformation defined by (7.12) preserves positive definiteness and symmetry of  $P(x)$  in  $P(q)$ , where the positive definiteness of  $P(x)$  is guaranteed by controllability and observability of (6.2) for all  $x$  [76]. Substituting (7.13) into (7.6) we thus find

$$\dot{V} = \frac{1}{2}q^T [A(q) + A^T(q) - 2P(q)]q = q^T \left[ \frac{A(q) + A^T(q)}{2} \right] q - q^T P(q)q \quad (7.14)$$

Now, using Rayleigh's quotient [57] we have

$$q^T \left[ \frac{A(q) + A^T(q)}{2} \right] q \leq \bar{\lambda} \left[ \frac{A(q) + A^T(q)}{2} \right] q^T q \quad (7.15)$$

$$q^T P(q) q \geq \underline{\lambda}[P(q)] q^T q \quad (7.16)$$

where  $\bar{\lambda}$  and  $\underline{\lambda}$  denote the maximum and minimum eigenvalues of a matrix, respectively. From (7.16)

we have  $-q^T P(q) q \leq -\underline{\lambda}[P(q)] q^T q$  so that we may write

$$\dot{V} \leq \left( \bar{\lambda} \left[ \frac{A(q) + A^T(q)}{2} \right] - \underline{\lambda}[P(q)] \right) q^T q \quad (7.17)$$

Finally, recognizing that  $\bar{\lambda} \left[ \frac{A+A^T}{2} \right]$  is the matrix measure of  $A$  with respect to the vector (Euclidean) 2-norm [69], we define it as

$$\mu_2[A(q)] \equiv \bar{\lambda} \left[ \frac{A(q) + A^T(q)}{2} \right] \quad (7.18)$$

so that (7.17) becomes

$$\dot{V} \leq (\mu_2[A(q)] - \underline{\lambda}[P(q)]) q^T q \quad (7.19)$$

We now seek to prove that the right hand side of (7.19) is always negative, and to do so we turn to properties of the algebraic Riccati equation. Observe that with  $A(x) = \bar{B}A(q)\bar{B}^{-1}$  and  $BR^{-1}B^T = \bar{B}\bar{B}^T$ , the continuous time SDRE (7.2) becomes

$$[\bar{B}A(q)\bar{B}^{-1}]^T P(x) + P(x)[\bar{B}A(q)\bar{B}^{-1}] - P(x)\bar{B}\bar{B}^T P(x) + Q(x) = 0 \quad (7.20)$$

Thus, premultiplying by  $\bar{B}^T$  and postmultiplying by  $\bar{B}$ , (7.20) becomes

$$A^T(q)\bar{B}^T P(x)\bar{B} + \bar{B}^T P(x)\bar{B}A(q) - \bar{B}^T P(x)\bar{B}\bar{B}^T P(x)\bar{B} + \bar{B}^T Q(x)\bar{B} = 0 \quad (7.21)$$

so that by defining

$$Q(q) \equiv \bar{B}^T Q(x)\bar{B} \quad (7.22)$$

and recalling (7.11) and (7.12) we may write (7.21) as

$$A^T(q)P(q) + P(q)A(q) - P(q)P(q) + Q(q) = 0 \quad (7.23)$$

Note that with  $H^T(x)H(x) = Q(x) > 0$  and  $\bar{B}$  full rank we have  $Q(q) > 0 \forall q$ . Now, we make use of a lower bound on the minimum eigenvalue of the solution of the continuous time algebraic Riccati equation

$$A^T P + PA - PBB^T P + Q = 0 \quad (7.24)$$

given in [49], which is

$$\underline{\lambda}[P] \geq \frac{-\mu_A + \sqrt{\mu_A^2 + \bar{\lambda}(BB^T)\mu_A^2\lambda\{(A^T A)^{-1}Q\}}}{\bar{\lambda}(BB^T)} \quad (7.25)$$

where

$$\mu_A \equiv \max\{-\mu_2(-A), -\mu_2(A)\} \quad (7.26)$$

and  $\mu_2(A)$  is as defined in (7.18). Note that for (7.25) to be valid, we must have  $\bar{\lambda}(BB^T) > 0$  and  $(A^T A)^{-1}$  must exist. Observe that from (7.23) we have  $B(q)B^T(q) = I$  so that clearly  $\bar{\lambda}[B(q)B^T(q)] = \underline{\lambda}[B(q)B^T(q)] = 1$ , and the first condition above is satisfied. As discussed above we also have  $A(q)$  nonsingular for all  $q$ , so that the second condition is also satisfied, and thus (7.25) may be applied to (7.23). Finally, we clearly have  $\mu_A(q) \geq -\mu_2(A(q))$  so that (7.25) gives

$$\underline{\lambda}[P(q)] \geq \mu_2[A(q)] + \sqrt{\mu_2^2[A(q)] + \mu_2^2[A(q)]\lambda\{[A^T(q)A(q)]^{-1}Q(q)\}} \quad (7.27)$$

Now, if  $\mu_2[A(q)] \neq 0$ , then clearly (7.27) implies

$$\underline{\lambda}[P(q)] > \mu_2[A(q)] \quad (7.28)$$

so that

$$\mu_2[A(q)] - \underline{\lambda}[P(q)] < 0 \quad (7.29)$$

and from (7.19) we see that  $\dot{V} < 0$ . If  $\mu_2[A(q)] = 0$ , then (7.19) becomes

$$\dot{V} \leq -\underline{\lambda}[P(q)]q^T q \quad (7.30)$$

and once again we see that  $\dot{V} < 0$ . Finally, the positive definiteness of  $P(q)$  for all  $q$  guarantees that  $q^T P(q)q$  vanishes only at the origin, and this completes the proof. ■

We note that in the above proof, it was not necessary that the control be applied to the system in transformed  $q$  coordinates. The control was applied in the original set of  $x$  coordinates, and only the stability analysis was performed in the  $q$  coordinates. One could, of course, first transform the system to have identity  $B$  matrix (including the effects of  $R$ ), and the above result would still hold. Indeed, this observation suggests that we could obtain a globally asymptotically stabilizing

SDRE regulator for systems with nonconstant but full rank  $B$  matrix functions for all  $x$ , by defining  $u(x) = B^{-1}(x)v(x)$ , and using the SDRE regulation algorithm to find the new control  $v$ .

In seeking to apply Theorem 7.3.1, we may find the nonsingularity constraint on  $A$  rather restrictive, and we thus seek a means of relaxing this requirement. This may be accomplished by defining

$$N_s \equiv \{q \in \mathcal{R}^n \mid \det[A(q)] = 0\} \quad (7.31)$$

i.e.,  $N_s$  is the set of all points in  $q$ -space where  $A(q)$  is singular, and making various assumptions about the nature of  $N_s$ . We now present several theorems in this vein.

**Theorem 7.3.2** *Assume all the conditions given in Theorem 7.3.1 hold except for  $a(x) = 0 \Rightarrow x = 0$ . Define  $N_s$  as in (7.31) and assume that for all  $q \in N_s$ , that  $A(q)q = \bar{B}^{-1}A(x)x = 0$ . Then the closed loop system defined by (6.2), (7.1), and (7.2) is globally asymptotically stable.*

**Proof:** With  $V$  defined as in (7.5), the proof that  $\dot{V}(q) < 0$  for all  $q$  not in  $N_s$  follows as per the proof of Theorem 7.3.1. For  $q \in N_s$ , since  $A(q)q = q^T A^T(q)$  vanishes, we have  $\dot{V} = -2q^T P(q)q < 0$  by the global positive definiteness of  $P(q)$ . Thus, we have a globally positive definite, decrescent, and radially unbounded Lyapunov function with globally negative derivative, so that the theorem is proven. ■

The point of the above theorem is that it is *not* the open loop equilibrium points that cause a potential problem for stability, since  $\dot{V} < 0$  at such points. It is instead the points at which  $A$  is singular but the state derivative does not vanish that are potentially troublesome for the proof. If we therefore define the set

$$N \equiv \{q \in \mathcal{R}^n \mid \det[A(q)] = 0 \text{ and } A(q)q \neq 0\} \quad (7.32)$$

it is on this set which we must make additional assumptions to guarantee stability, as we do in the following theorems.

**Theorem 7.3.3** *Assume all the conditions given in Theorem 7.3.1 hold except for  $a(x) = 0 \Rightarrow x = 0$ . Define  $N$  as in (7.32) and assume  $N$  is bounded. Then all trajectories of the closed loop system defined by (6.2), (7.1), and (7.2) are globally bounded.*

**Proof:** The conclusion follows trivially from the existence of a globally positive definite, decrescent, and radially unbounded Lyapunov function (7.5) which has negative derivative outside of a compact set containing the origin [40]. ■

If we now define the set

$$E \equiv \{q \in R^n \mid q^T [A(q) + A^T(q) - 2P(q)]q = 0\} \quad (7.33)$$

and make an assumption on  $E$  and an additional assumption on  $N$ , then we may strengthen Theorem 7.3.3 to conclude global asymptotic stability as presented in the following theorem.

**Theorem 7.3.4** *Assume all the conditions of Theorem 7.3.3 hold, and further assume that  $\dot{V} \leq 0$  on  $N$ , and that  $E$  contains only the single element  $q = 0$ . Then the closed loop system defined by (6.2), (7.1), and (7.2) is globally asymptotically stable.*

**Proof:** From the proofs of Theorems 7.3.1 and 7.3.3, we have that all trajectories of the closed loop system are globally bounded. Since the closed loop system is autonomous and  $\dot{V} \leq 0$  globally, Lasalle's invariance principle [44] implies that all trajectories converge to the maximal invariant set contained in  $E$ , which by assumption is the origin, and thus the theorem is proved. ■

We point out that the condition of Theorem 7.3.3 is not necessary for global asymptotic stability of the origin as proven in Theorem 7.3.4, if global boundedness of all the system trajectories may be established by some other means. It turns out that [49] provides two alternative lower bounds on the minimum eigenvalue of  $P$  of (7.24), which are

$$\lambda(P) \geq \frac{\frac{1}{2}\lambda\{(A + A^T)Q^{-1}\} + \sqrt{(\frac{1}{2}\lambda\{(A + A^T)Q^{-1}\})^2 + \bar{\lambda}(BB^TQ^{-1})}}{\bar{\lambda}(BB^TQ^{-1})} \quad (7.34)$$

and

$$\lambda(P) \geq \frac{-\bar{\sigma}(A) + \sqrt{\bar{\sigma}^2(A) + \bar{\lambda}(BB^T)\lambda(Q)}}{\bar{\lambda}(BB^T)} \quad (7.35)$$

These bounds may also be used to generate sufficient conditions for global asymptotic closed loop stability of the origin, as formalized in the following theorems.

**Theorem 7.3.5** *Assume all the conditions given in Theorem 7.3.1 hold except for  $a(x) = 0 \Rightarrow x = 0$ , and let  $N$  be as in (7.32). Also, assume that on some set  $L \subset \mathcal{R}^n$  we have  $Q(x) = k(\bar{B}\bar{B}^T)^{-1}$ , where  $k > 0$  is a real number, and assume that  $L \cup N = \mathcal{R}^n$ . Additionally, assume that for all  $q \in L$  we have either*

$$\mu_2[A(q)] < \underline{\lambda}[A(q) + A^T(q)] \text{ and } \underline{\lambda}[A(q) + A^T(q)] > 0 \quad (7.36)$$

or

$$k > \mu_2^2[A(q)] - \mu_2[A(q)]\underline{\lambda}[A(q) + A^T(q)] \quad (7.37)$$

*Then the closed loop system defined by (6.2), (7.1), and (7.2) is globally asymptotically stable.*

**Proof:** Choose the globally positive definite, decrescent, and radially unbounded Lyapunov function (7.5). From the proof of Theorem 7.3.1 we have  $\dot{V} < 0 \forall q \in N$ . Now, for all  $q \in L$ , we have  $Q(q) = kI$  from (7.22) and the stated assumptions. Recalling  $B(q)B^T(q) = I$  and substituting for  $Q(q)$  in (7.34) we find

$$\underline{\lambda}[P(q)] \geq \frac{1}{2}\underline{\lambda}[A(q) + A^T(q)] + \sqrt{\left(\frac{1}{2}\underline{\lambda}[A(q) + A^T(q)]\right)^2 + k} \quad (7.38)$$

and it is easily established that both (7.36) and (7.37) guarantee that  $\underline{\lambda}[P(q)] > \mu_2[A(q)]$ , so that from (7.19) we clearly have  $\dot{V} < 0 \forall q \in L$ . The assumption that  $L \cup N = \mathcal{R}^n$  then ensures that  $\dot{V} < 0 \forall q \in \mathcal{R}^n$ , which completes the proof. ■

We note that the above theorem allows the nonsingularity requirement on  $A(q)$  to be relaxed, but at the expense of having to choose a prescribed form of the state weighting matrix, with only the scalar  $k$  left as a design parameter, which must be chosen sufficiently large. The next theorem also allows relaxation of the nonsingularity requirement on  $A(q)$ , with the only requirement being that  $Q(q)$  is chosen ‘big enough.’ At this point it is interesting to note that Theorems 7.3.5 and 7.3.6 provide theoretical corroboration for a phenomenon observed in [19], namely, that state weightings

must sometimes be increased beyond initial choices to sufficiently large levels in order to provide global closed loop stability of an SDRE regulator.

**Theorem 7.3.6** *Assume all the conditions given in Theorem 7.3.1 hold except for  $a(x) = 0 \Rightarrow x = 0$ , and let  $N$  be as in (7.32). Also, assume that on some set  $M \subset \mathcal{R}^n$  we have for all  $q \in M$*

$$\Delta[Q(q)] > 2\bar{\sigma}[A(q)]\mu_2[A(q)] + \mu_2^2[A(q)] \quad (7.39)$$

*and assume that  $M \cup N = \mathcal{R}^n$ . Then the closed loop system defined by (6.2), (7.1), and (7.2) is globally asymptotically stable.*

**Proof:** The proof is exactly like that of Theorem 7.3.5, except (7.39) substituted into (7.35) (with  $B(q)B^T(q) = I$ ) guarantees that  $\dot{V} < 0 \forall q \in M$ . ■

Since the induced norm is an upper bound for the matrix measure [69], we have the following simplified condition which can be used in place of (7.39).

**Theorem 7.3.7** *Assume all the conditions of Theorem 7.3.6 hold, except that in place of (7.39) we have*

$$\Delta[Q(q)] > 3\bar{\sigma}^2[A(q)] \quad \forall q \in M \quad (7.40)$$

*Then the closed loop system defined by (6.2), (7.1), and (7.2) is globally asymptotically stable.*

**Proof:** The claim follows trivially from the proof of Theorem 7.3.6 and the fact that  $\bar{\sigma}[A] \geq \mu_2(A)$  for any  $A \in \mathcal{C}^n$ . ■

Finally, we observe that the above theorems can be combined in various forms, e.g., use Theorems 7.3.1, 7.3.5, and 7.3.6 with the assumption that  $L \cup M \cup N = \mathcal{R}^n$  in place of the similar assumptions therein.

## 7.4 Discussion

In this section we include some remarks concerning the full rank, constant  $B$  matrix assumption. Besides allowing the linear, static coordinate transformation (7.3), and use of the bound (7.25), the

full rank, constant  $B$  matrix assumption serves another purpose. In Chapter 6 it was shown that systems of the form (6.1) which may be further written as (6.2), with full rank  $B$  matrices, are nonlinearly controllable for all  $x$  in addition to having  $\{A(x), B\}$  controllable in the linear sense for all  $x$ , regardless of the choice of  $A$ . Thus, the reachable set from each  $x$  has dimension  $n$ , so that the uncontrollable subspace for such systems consists entirely of the zero vector. This is important because the proofs in this chapter guarantee that the bracketed term in the Lyapunov function derivative  $\dot{V} = q^T[\frac{A(q)+A^T(q)}{2} - P(q)]q$  is negative definite. However,  $-P(q)q = B(q)u(q)$ , so that if  $q$  is in the uncontrollable subspace for the system (6.1), then  $q^TB(q)u(q) = 0$  and we are left with  $\dot{V} = q^T[\frac{A(q)+A^T(q)}{2}]q$ , which we clearly cannot guarantee to be negative for all  $q$  except under extremely restrictive conditions. The full rank  $B$  matrix assumption thus guarantees that  $q^TB(q)u(q) = 0$  only for  $q = 0$ . Finally, notice that we may allow  $u \in \mathcal{R}^m$ , where  $m \geq n$ , and  $\text{rank}(B) = n$ . In this case, just replace  $\bar{B}^{-1}$  with  $\bar{B}^\dagger$ , where  $(\cdot)^\dagger$  denotes the pseudoinverse.

## 7.5 Conclusion

We have given conditions which guarantee global asymptotic stability of a class of nonlinear, input-affine systems employing SDRE nonlinear regulation. The conditions common to every theorem are: full rank, constant  $B$  matrices; constant, positive definite control penalty matrix  $R$ ; and positive definite state weighting matrix function  $Q(x)$  for all  $x$ . Various additional assumptions on the choice of  $Q$  or  $A$  are combined to obtain the final sufficient stability conditions. The full rank  $B$  matrix assumption was also said to be important in that it guarantees an uncontrollable subspace consisting of only the zero vector, thus validating the expressions establishing negativity of the Lyapunov function derivative which rely on negative definiteness of an  $n \times n$  matrix at each point in the state space.

# VIII. Lyapunov Stability of Analytic Sampled Data Systems with Positive Definite $Q$ Matrices

## 8.1 Introduction

In this chapter we consider sampled data implementation of the continuous SDRE nonlinear regulation algorithm, and derive sufficient conditions for theoretically global asymptotic stability of the closed loop system. The global theoretical result obtained requires infinitely fast sampling, so that real applications will have bounded domains of attraction, and thus it is perhaps more accurate to say semiglobal stability for the sampled data regulator is proven.

We again consider the continuous time, state feedback, input-affine, autonomous nonlinear dynamic system described by

$$\begin{aligned} \dot{x} &= a(x) + b(x)u, \quad a(0) = 0 \\ z &= \begin{bmatrix} h(x) \\ \bar{R}(x)u \end{bmatrix}, \quad h(0) = 0 \end{aligned} \tag{8.1}$$

with state vector  $x \in \mathcal{R}^n$ , control  $u \in \mathcal{R}^m$ , penalized variable  $z \in \mathcal{R}^s$ , and control penalty matrix  $\bar{R}(x)$  (assumed nonsingular for all  $x$ ). We assume that the  $a_i$ ,  $B_{ij}$ ,  $h_i$  and  $\bar{R}_{ij}$  are all real-valued analytic functions of  $x$  on  $\mathcal{R}^n$ . As shown in Section 2.5, under the above assumptions the system (8.1) can be written [69] (nonuniquely) in the state-dependent coefficient (SDC) form

$$\begin{aligned} \dot{x} &= A(x)x + B(x)u \\ z &= \begin{bmatrix} H(x)x \\ \bar{R}(x)u \end{bmatrix} \end{aligned} \tag{8.2}$$

where  $A$  and  $H$  are chosen to be analytic in  $x$  (see Section 2.5.2). We also assume that the pairs  $\{A(x), B(x)\}$  and  $\{H(x), A(x)\}$  are globally stabilizable and detectable, respectively. Furthermore, we assume that  $H(x)$  is globally nonsingular so that  $H^T(x)H(x)$  has full rank for all  $x$ . Note that this assumption on  $H$  actually strengthens the global detectability assumption on  $\{H(x), A(x)\}$

to global observability. Using these assumptions we will now establish sufficient conditions under which the sampled data state feedback nonlinear SDRE regulator yields a closed loop system which is theoretically globally asymptotically stable.

## 8.2 Sampled Data Nonlinear SDRE Regulation

The continuous time control algorithm is given as before by

$$u = -R^{-1}(x)B^T(x)P(x)x \quad (8.3)$$

where  $P(x)$  is the maximal solution to the steady state continuous time state-dependent Riccati equation (SDRE)

$$A^T(x)P(x) + P(x)A(x) - P(x)B(x)R^{-1}(x)B^T(x)P(x) + H^T(x)H(x) = 0 \quad (8.4)$$

which is guaranteed to exist under the stabilizability/detectability assumptions, where we have defined  $R(x) = \bar{R}^T(x)\bar{R}(x) > 0$ . In actual implementation, we will use a sampled data form of the above, taking measurements (assumed perfect) of the state variables periodically and using them to generate new values of the control. Adopting the convention  $x(t_k) = x_k$  and  $A(x(t_k)) = A_k$  (and likewise for all other vectors and matrices), the algorithm is to apply the control

$$u_k = -R_k^{-1}B_k^T P_k x_k \quad (8.5)$$

at each sampling time  $t_k$ , where  $P_k$  is obtained from the sampled data SDRE

$$A_k^T P_k + P_k A_k - P_k B_k R_k^{-1} B_k^T P_k + H_k^T H_k = 0 \quad (8.6)$$

Note that we are applying a constant control input over any given sampling interval  $[t_k, t_{k+1})$ , and we have left the sampling interval size,  $\delta_k = t_{k+1} - t_k$ , unspecified as of yet. We emphasize that the above algorithm is simply a sampled data implementation of the continuous time SDRE regulator, and not a discrete time control law derived from a discrete time dynamical system model. The motivation for the above approach is to gain a better understanding of the continuous controller through a discrete time approximation. A more rigorous discrete time treatment of the SDRE

regulator problem is not pursued in this dissertation, but is instead proposed as a topic for further research. To prove stability of this algorithm, we shall use the following discrete time version of the Lyapunov stability theorem, taken from [7].

**Theorem 8.2.1** *Consider the vector difference equation*

$$x_{k+1} = \Phi_k x_k \quad (8.7)$$

*with transition matrix*

$$\Phi(t_{k+N}, t_k) = \Phi_{k+N-1} \Phi_{k+N-2} \dots \Phi_{k+1} \Phi_k. \quad (8.8)$$

*Suppose there exist  $\alpha, \beta \in \mathbb{R}$  and a positive definite matrix sequence  $X_k$  with  $0 < \alpha I \leq X_k \leq \beta I < \infty$  such that*

$$\Phi_k^T X_{k+1} \Phi_k - X_k = -N_k^T N_k \quad (8.9)$$

*for some matrix sequence  $N_k$ , and all  $k$ . Then (8.7) is stable in the sense of Lyapunov, and  $V_k = x_k^T X_k x_k$  is a Lyapunov function for (8.7). Furthermore, if  $N_k > 0$  for all  $k$ , then (8.7) is asymptotically stable. Alternatively, if  $N_k \geq 0$  and  $\Delta V_k = -x_k^T N_k^T N_k x_k$  does not vanish identically for all  $k$ , any  $t_0$ , and any  $x_0$ , then (8.7) is asymptotically stable.*

**Proof:** See [6] and [36]. ■

### 8.3 Linear Transition Matrix Representation

We first show that the closed loop system using (8.5) can be written in the form of (8.7), provided  $\delta_k$  is chosen small enough. Using (8.2), for  $t \in [t_k, t_{k+1})$  the closed loop dynamics become

$$\dot{x} = A(x)x + B(x)u_k \equiv f(x, x_k) \quad (8.10)$$

Recall that by assumption  $a(x) = A(x)x$  and  $B(x)$  are analytic with respect to  $x$  on  $\mathcal{R}^n$ , and so is  $u_k$  since it is just a constant vector. Thus,  $f$  is analytic over any sampling interval, implying it has a convergent power series representation in  $x$  and thus that it has continuously differentiable partial derivatives. This last observation is sufficient to guarantee that  $f$  satisfies a local Lipschitz condition

for any  $x_k \in \mathcal{R}^n$  [1], so that (8.10) possesses a unique solution near  $x_k$  which is continuous and in fact analytic in  $t$  if the system matrices are analytic in  $x$  [36]. Thus, by taking  $t_{k+1}$  sufficiently close to  $t_k$ , we can ensure that  $x$  is continuous (and analytic) with respect to  $t$ . Continuity of  $x$  ensures continuity of  $f$  with respect to  $t$  over the sampling interval, and thus we can integrate (8.10) from  $t_k$  to  $t_{k+1}$  to obtain

$$x_{k+1} = x_k + \int_{t_k}^{t_{k+1}} f(x(t), x_k) dt \quad (8.11)$$

$$= x_k + \int_{t_k}^{t_{k+1}} A(x(t))x(t) dt - \left( \int_{t_k}^{t_{k+1}} B(x(t)) dt \right) R_k^{-1} B_k^T P_k x_k \quad (8.12)$$

Now we can apply the Mean Value Theorem (MVT) [1] for integrals to each individual element of the right hand side of (8.11) to obtain

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{k+1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_k + \delta_k \begin{bmatrix} f_1(t_k^{(1)}) \\ f_2(t_k^{(2)}) \\ \vdots \\ f_n(t_k^{(n)}) \end{bmatrix} \quad (8.13)$$

where  $\delta_k = t_{k+1} - t_k$  and  $t_k^{(i)} \in [t_k, t_{k+1}] \forall i \in [1, n]$  is the time of evaluation of  $f_i$  in the MVT.

We have introduced a slight abuse of notation by letting  $f_i$  represent the composite function of  $t$  as well as the function of  $x$ . The notation on  $t$  is as follows: the superscripted number in parenthesis indicates the row of  $f$  that is being evaluated using the MVT, while the subscript  $k$  simply indicates association with the sampling interval start time  $t_k$ . Using the definition of  $f$  we get

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{k+1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_k + \delta_k \begin{bmatrix} \Sigma_j A_{1j}(t_k^{(1)})x_j(t_k^{(1)}) \\ \Sigma_j A_{2j}(t_k^{(2)})x_j(t_k^{(2)}) \\ \vdots \\ \Sigma_j A_{nj}(t_k^{(n)})x_j(t_k^{(n)}) \end{bmatrix} - \delta_k \bar{B}_k R_k^{-1} B_k^T P_k x_k \quad (8.14)$$

where we have defined

$$\bar{B}_k \equiv \begin{bmatrix} B_{11}(t_k^{(1)}) & B_{12}(t_k^{(1)}) & \cdots & B_{1m}(t_k^{(1)}) \\ B_{21}(t_k^{(2)}) & B_{22}(t_k^{(2)}) & \cdots & B_{2m}(t_k^{(2)}) \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1}(t_k^{(n)}) & B_{n2}(t_k^{(n)}) & \cdots & B_{nm}(t_k^{(n)}) \end{bmatrix} \quad (8.15)$$

Consider now only the second term on the right hand side of (8.14). This term may be decomposed

into a sum of  $n$  terms as

$$\begin{aligned} & \delta_k \begin{bmatrix} A_{11}(t_k^{(1)}) & A_{12}(t_k^{(1)}) & \cdots & A_{1n}(t_k^{(1)}) \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} x(t_k^{(1)}) \\ & + \delta_k \begin{bmatrix} 0 & 0 & \cdots & 0 \\ A_{21}(t_k^{(2)}) & A_{22}(t_k^{(2)}) & \cdots & A_{2n}(t_k^{(2)}) \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} x(t_k^{(2)}) + \cdots \\ & + \delta_k \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ A_{n1}(t_k^{(n)}) & A_{n2}(t_k^{(n)}) & \cdots & A_{nn}(t_k^{(n)}) \end{bmatrix} x(t_k^{(n)}) \quad (8.16) \\ & = \delta_k [\bar{A}_k^{(1)} x(t_k^{(1)}) + \bar{A}_k^{(2)} x(t_k^{(2)}) + \cdots + \bar{A}_k^{(n)} x(t_k^{(n)})] \quad (8.17) \end{aligned}$$

Note that  $\bar{A}_k^{(i)}$  is a matrix with its  $i$ th row equal to the  $i$ th row of  $A$  evaluated at  $t_k^{(i)}$ , and zero elsewhere. Now, by taking advantage of the fact that each of the  $x(t_k^{(i)})$  above may in turn be written as

$$x(t_k^{(i)}) = x_k + \int_{t_k}^{t_k^{(i)}} f(x(t), x_k) dt \quad (8.18)$$

and again applying the MVT to the integrals in (8.18), we see that (8.17) is equal to

$$\sum_{i=1}^n \delta_k \bar{A}_k^{(i)} \left( x_k + \delta_k^{(i)} \begin{bmatrix} \Sigma_j A_{1j}(t_k^{(i1)}) x_j(t_k^{(i1)}) \\ \Sigma_j A_{2j}(t_k^{(i2)}) x_j(t_k^{(i2)}) \\ \vdots \\ \Sigma_j A_{nj}(t_k^{(in)}) x_j(t_k^{(in)}) \end{bmatrix} - \delta_k^{(i)} \bar{B}_k^{(i)} R_k^{-1} B_k^T P_k x_k \right) \quad (8.19)$$

where  $\delta_k^{(i)} \equiv t_k^{(i)} - t_k$  and

$$\bar{B}_k^{(i)} \equiv \begin{bmatrix} B_{11}(t_k^{(i1)}) & B_{12}(t_k^{(i1)}) & \cdots & B_{1m}(t_k^{(i1)}) \\ B_{21}(t_k^{(i2)}) & B_{22}(t_k^{(i2)}) & \cdots & B_{2m}(t_k^{(i2)}) \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1}(t_k^{(in)}) & B_{n2}(t_k^{(in)}) & \cdots & B_{nm}(t_k^{(in)}) \end{bmatrix} \quad (8.20)$$

The notation  $t_k^{(ij)} \in [t_k, t_k^{(i)}]$  indicates this time is generated from the  $i$ th previous value of  $t_k$ , and that the  $j$ th row of  $f$  is being evaluated using the MVT. The subscript  $k$  is again to indicate dependence on the original starting time. Note we have  $t_k^{(i)} \leq t_{k+1} \forall i$  so that all  $\delta_k^{(i)} \leq \delta_k$ . We can continue with another application of the MVT, noting that each application generates  $n$  new  $t$  values, and we adapt our notation by adding an additional superscript to  $t_k$ ,  $\delta_k$ ,  $\bar{A}_k$  and  $\bar{B}_k$ . By defining

$$\bar{A}_k^{(\Sigma)} = \sum_{i=1}^n \bar{A}_k^{(i)} \quad (8.21)$$

and defining  $\bar{A}_k^{(ij)}$  for (8.19) analogously to how  $\bar{A}_k^{(i)}$  was defined in (8.17), and subsequently setting

$$\bar{A}_k^{(i\Sigma)} = \sum_{j=1}^n \bar{A}_k^{(ij)}, \quad (8.22)$$

then (8.19) becomes

$$\delta_k \left[ \bar{A}_k^{(\Sigma)} + \sum_{i=1}^n \delta_k^{(i)} \bar{A}_k^{(i)} \bar{A}_k^{(i\Sigma)} - \sum_{i=1}^n \delta_k^{(i)} \bar{B}_k^{(i)} R_k^{-1} B_k^T P_k \right] x_k + \mathcal{O}(\delta^3, \bar{A}^3) x(t_k^{(ijl)}) \quad (8.23)$$

where the  $\mathcal{O}$  notation indicates terms of order 3 in both  $\delta$  and  $\bar{A}$  premultiplying the  $n^3$  terms in  $x$ , and  $l$  is another integer index ranging from 1 to  $n$ . Now, (8.18) and the MVT may be repeatedly applied to the  $x(t_k^{(ijl\cdots)})$  of (8.23), leading to an infinite matrix series expression in  $\delta$  and  $\bar{A}$ , which, if

convergent, implies that the higher-order terms contribute nothing (the higher-order terms are the limit of the sequence of  $\delta \bar{A}$  products as the number of terms tends to infinity, which must be zero for a convergent series). Furthermore, each time the MVT is applied to an expression of the form of (8.18), it introduces another term with a single  $\bar{B}_k^{(\cdot)}$  postmultiplier in it – thus, another infinite series. Assuming both these series converge, then (8.11) may be written

$$x_{k+1} = \left[ I + \delta_k \bar{A}_k^{(\Sigma)} + \delta_k \sum_{i=1}^n \delta_k^{(i)} \bar{A}_k^{(i)} \bar{A}_k^{(i\Sigma)} + \dots \right] x_k - \delta_k \left[ \bar{B}_k + \sum_{i=1}^n \delta_k^{(i)} \bar{A}_k^{(i)} \bar{B}_k^{(i)} + \dots \right] R_k^{-1} B_k^T P_k x_k \quad (8.24)$$

This yields a transition matrix representation, as formally stated and proven in the following theorem.

**Theorem 8.3.1** *Consider the continuous time open loop system (8.2) with discrete time control law given by (8.5) and (8.6), where  $A(x), B(x), H(x)$  and  $R(x)$  are analytic in  $x \in \mathcal{R}^n$ , and that the pairs  $\{A(x), B(x)\}$  and  $\{H(x), A(x)\}$  are stabilizable and detectable, respectively, for all  $x$ . Define*

$$B_k(A) = \max_{i,j,t \in [t_k, t_{k+1}]} |A_{ij}(t)| \quad (8.25)$$

*Then if the sampling interval size  $\delta_k = t_{k+1} - t_k$  is chosen such that*

$$\delta_k < \frac{1}{n B_k(A)} \quad (8.26)$$

*the closed loop system (8.10) can be written in the form*

$$x_{k+1} = \Phi_k x_k \quad (8.27)$$

*where*

$$\Phi_k = \left[ I + \delta_k \bar{A}_k^{(\Sigma)} + \delta_k \sum_{i=1}^n \delta_k^{(i)} \bar{A}_k^{(i)} \bar{A}_k^{(i\Sigma)} + \dots \right] - \delta_k \left[ \bar{B}_k + \sum_{i=1}^n \delta_k^{(i)} \bar{A}_k^{(i)} \bar{B}_k^{(i)} + \dots \right] R_k^{-1} B_k^T P_k \quad (8.28)$$

**Proof:** By the development involving the MVT in (8.11) - (8.24), we must simply prove that both infinite matrix series are convergent for (8.28) to be valid. Since a matrix series converges iff each individual matrix element series converges [42], we need to prove element-wise convergence.

Recall that the  $A_{ij}$  and  $B_{ij}$  are continuous functions of  $x$ , and we have restricted  $t_{k+1}$  so that  $x$  is a continuous function of  $t$  on the closed, bounded interval  $[t_k, t_{k+1}]$ . Thus, the  $A_{ij}$  and  $B_{ij}$ , and in fact the absolute values  $|A_{ij}|$  and  $|B_{ij}|$ , are continuous functions of  $t$  on a compact subset of  $\mathcal{R}$ , so that each  $|A_{ij}|$  and  $|B_{ij}|$  attains a maximum over  $[t_k, t_{k+1}]$ . Let  $\mathcal{B}_k(A)$  be defined as in (8.25) and

$$\mathcal{B}_k(B) = \max_{i,j,t \in [t_k, t_{k+1}]} |B_{ij}(t)| \quad (8.29)$$

From (8.28), define

$$\Psi_k = I + \delta_k \bar{A}_k^{(\Sigma)} + \delta_k \sum_{i=1}^n \delta_k^{(i)} \bar{A}_k^{(i)} \bar{A}_k^{(i\Sigma)} + \dots \quad (8.30)$$

$$\Gamma_k = \delta_k \bar{B}_k + \delta_k \sum_{i=1}^n \delta_k^{(i)} \bar{A}_k^{(i)} \bar{B}_k^{(i)} + \dots \quad (8.31)$$

with the higher-order terms determined by repeated application of the MVT as described above.

Then from the definitions in (8.25) and (8.29) and the fact that  $\delta_k \geq \delta_k^{(\cdot)}$ , we have

$$\begin{aligned} |\Psi_{kij}| &\leq 1 + \delta_k \mathcal{B}_k(A) + \delta_k^2 n \mathcal{B}_k^2(A) + \delta_k^3 n^2 \mathcal{B}_k^3(A) + \dots \\ &= 1 - \frac{1}{n} + \frac{1}{n} \left\{ \sum_{r=0}^{\infty} [\delta_k n \mathcal{B}_k(A)]^r \right\} \end{aligned} \quad (8.32)$$

and

$$\begin{aligned} |\Gamma_{kij}| &\leq \delta_k \mathcal{B}_k(B) + \delta_k^2 n \mathcal{B}_k(A) \mathcal{B}_k(B) + \delta_k^3 n^2 \mathcal{B}_k^2(A) \mathcal{B}_k(B) + \dots \\ &= \delta_k \left\{ \sum_{r=0}^{\infty} [\delta_k n \mathcal{B}_k(A)]^r \right\} \mathcal{B}_k(B) \end{aligned} \quad (8.33)$$

Therefore, both  $|\Psi_{kij}|$  and  $|\Gamma_{kij}|$  will be convergent so long as the geometric series

$$\sum_{r=0}^{\infty} [\delta_k n \mathcal{B}_k(A)]^r \quad (8.34)$$

converges. It is well known that a sufficient condition for convergence of (8.34) is to require

$$|\delta_k n \mathcal{B}_k(A)| < 1 \quad (8.35)$$

Since absolute convergence of an infinite series implies convergence of the series [1], we may ensure validity of (8.28) by choosing  $\delta_k$  according to (8.26).

We have thus proven that our system dynamics can be written in the form of (8.27), provided each  $\delta_k$  is chosen small enough. Note that if  $\delta_k$  is chosen small enough to guarantee validity of (8.28) up to the next sampling time  $t_{k+1}$ , then clearly we also have such a valid expression for any  $x$  within the sampling interval, i.e., any  $x_r$  such that  $t_r \in [t_k, t_{k+1}]$ . ■

#### 8.4 Stability Via Lyapunov Theory

We shall now make use of Theorem 8.2.1 to prove stability of the sampled data discrete time system, which we now write via Theorem 8.3.1 as

$$x_{k+1} = \Phi_k x_k = [I + \delta_k(\bar{A}_k^{(\Sigma)} - \bar{B}_k R_k^{-1} B_k^T P_k) + \mathcal{O}(\delta^2, \bar{A}^2) + \mathcal{O}(\delta^2, \bar{A}\bar{B})]x_k \quad (8.36)$$

where the last two terms in the bracketed expression represent higher-order terms in the infinite series expansion of  $\Phi_k$ . We first consider only variations over the entire sampling interval, and then address intersample behavior. Define

$$\Delta A_k \equiv \bar{A}_k^{(\Sigma)} - A_k \quad (8.37)$$

$$\Delta B_k \equiv \bar{B}_k - B_k \quad (8.38)$$

Then (8.36) may be written

$$\begin{aligned} x_{k+1} = & [I + \delta_k(A_k - B_k R_k^{-1} B_k^T P_k + \Delta A_k - \Delta B_k R_k^{-1} B_k^T P_k) \\ & + \mathcal{O}(\delta^2, \bar{A}^2) + \mathcal{O}(\delta^2, \bar{A}\bar{B})]x_k \end{aligned} \quad (8.39)$$

Now, recognizing that  $A_k - B_k R_k^{-1} B_k^T P_k$  is just the closed loop continuous dynamics matrix sampled at time  $t_k$ , we define this matrix as  $F_k$  to get

$$x_{k+1} = [I + \delta_k F_k + \delta_k(\Delta A_k - \Delta B_k R_k^{-1} B_k^T P_k) + \mathcal{O}(\delta^2, \bar{A}^2) + \mathcal{O}(\delta^2, \bar{A}\bar{B})]x_k \quad (8.40)$$

Recalling Theorem 8.2.1, we seek to prove that there exists a bounded positive definite matrix sequence  $X_k$  such that the Lyapunov function  $V_k = x_k^T X_k x_k$  has a negative variation  $\Delta V_k$  along the trajectory for all  $k$ . Using (8.40) we see that

$$\begin{aligned}
\Delta V_k &= V_{k+1} - V_k \\
&= x_{k+1}^T X_{k+1} x_{k+1} - x_k^T X_k x_k \\
&= x_k^T [\Phi_k^T X_{k+1} \Phi_k - X_k] x_k \\
&= x_k^T \left[ X_{k+1} - X_k + \delta_k \left\{ F_k^T X_{k+1} + X_{k+1} F_k + \Delta A_k^T X_{k+1} + X_{k+1} \Delta A_k \right. \right. \\
&\quad \left. \left. - (\Delta B_k R_k^{-1} B_k^T P_k)^T X_{k+1} - X_{k+1} (\Delta B_k R_k^{-1} B_k^T P_k) \right\} \right. \\
&\quad \left. + \mathcal{O}(\delta^2, A^{2-m} B^m) \right] x_k
\end{aligned} \tag{8.41}$$

where  $m = \{0, 1, 2\}$  and the  $\mathcal{O}$  notation represents terms of order 2 or higher in  $\delta$ , which contain  $A$  and  $B$  terms of a combined order no greater than the order of  $\delta$ . We now prove that under certain conditions there always exists such a sequence  $X_k$ . The proof is constructive by nature; i.e., we show how such a sequence may be constructed. Before proceeding we make a crucial observation concerning the nature of the solution to the continuous time SDRE (8.4). It has been proven [7] that, if the elements of the matrices  $A$ ,  $B$ ,  $R$ ,  $H$  involved in a continuous time algebraic Riccati equation depend continuously on some parameter and if stabilizability and detectability assumptions hold, then the maximal solution  $P$  also depends continuously on that parameter. In fact, it has been shown [58] that the same may be said for analyticity with respect to a parameter under the assumptions of analyticity of the matrix elements and stabilizability and detectability.

**Theorem 8.4.1** *Assume the conditions in Theorem 8.3.1 hold, and further assume that  $H(x)$  is globally nonsingular so that  $H^T(x)H(x) > 0$  for all  $x$ , and that the solution to the SDRE,  $P_k$ , is such that  $\lim_{k \rightarrow \infty} P_k$  exists. Then for  $\delta_k$  sufficiently small, the conditions of Theorem 8.2.1 hold and the closed loop system (8.10) is globally asymptotically stable.*

**Proof:** First, note that the SDRE solution  $P(x)$  is positive definite for any  $x$ , since  $\{H(x), A(x)\}$  is observable under our assumption of global nonsingularity of  $H(x)$  [76]. For any  $t_r \in (t_k, t_{k+1}]$ ,

define

$$\begin{aligned}\Delta_r &= \frac{x_k^T(P_r - P_k)x_k}{\delta_r x_k^T x_k} \quad (\text{if } x_k \neq 0) \\ &= 0 \quad (\text{if } x_k = 0)\end{aligned}\tag{8.42}$$

where  $\delta_r = t_r - t_k$  and let

$$X_r = s_r P_r\tag{8.43}$$

where  $P_r$  is the SDRE solution at  $x_r$ , and  $s_r$  is a positive scalar defined by the recursive expression

$$\begin{aligned}s_r &= \frac{x_k^T P_k x_k}{x_k^T P_r x_k} s_k \quad (\text{if } \Delta_r \geq \epsilon) \\ &= s_k \quad (\text{otherwise})\end{aligned}\tag{8.44}$$

with  $s_0 = 1$  and where  $\epsilon$  is an arbitrarily small positive tolerance selected so that

$$\epsilon < \min_k [\underline{\lambda}(H_k^T H_k)]\tag{8.45}$$

where  $\underline{\lambda}$  is the minimum eigenvalue. Using (8.43), (8.41) becomes

$$\begin{aligned}\Delta V_k &= x_k^T [s_{k+1} P_{k+1} - s_k P_k] x_k \\ &\quad + s_{k+1} \delta_k x_k^T \left[ F_k^T P_{k+1} + P_{k+1} F_k + \Delta A_k^T P_{k+1} + P_{k+1} \Delta A_k \right. \\ &\quad \left. - (\Delta B_k R_k^{-1} B_k^T P_k)^T P_{k+1} - P_{k+1} (\Delta B_k R_k^{-1} B_k^T P_k) \right] x_k \\ &\quad + x_k^T \mathcal{O}(\delta^2, A^{2-m} B^m) x_k\end{aligned}\tag{8.46}$$

Notice that if  $N_k > 0$  in (8.9), then (8.9) is equivalent to (8.41) being negative for all  $x_k$  other than zero, and thus we want (8.46) to be negative for all  $x_k$  other than zero. From (8.42) and (8.44) we have

$$x_k^T [s_{k+1} P_{k+1} - s_k P_k] x_k \leq \epsilon s_{k+1} \delta_k x_k^T x_k \quad \forall k\tag{8.47}$$

so that the first term in (8.46) is always bounded by an arbitrarily small positive constant. We now make use of the analyticity of the  $P_{ij}$ ,  $A_{ij}$  and  $B_{ij}$  with respect to  $t$  to expand in a neighborhood

of  $t_k$  and write

$$P_{ij_r} = P_{ij_k} + \frac{\partial P_{ij}}{\partial t} \bigg|_{t_k} \delta_r + \dots = P_{ij_k} + P'_{ij_k} \delta_r + \dots \quad (8.48)$$

$$A_{ij_r} = A_{ij_k} + \frac{\partial A_{ij}}{\partial t} \bigg|_{t_k} \delta_r + \dots = A_{ij_k} + A'_{ij_k} \delta_r + \dots \quad (8.49)$$

$$B_{ij_r} = B_{ij_k} + \frac{\partial B_{ij}}{\partial t} \bigg|_{t_k} \delta_r + \dots = B_{ij_k} + B'_{ij_k} \delta_r + \dots \quad (8.50)$$

for  $t_r \in [t_k, t_{k+1}]$ , where  $P'_k$ ,  $A'_k$  and  $B'_k$  are the matrices of partials of  $P$ ,  $A$ , and  $B$  with respect to  $t$  evaluated at  $t_k$ , respectively. Using (8.37), (8.38), and (8.48), (8.46) becomes

$$\begin{aligned} \Delta V_k \leq & s_{k+1} \delta_k x_k^T \left[ \epsilon I + F_k^T (P_k + P'_k \delta_k) + (P_k + P'_k \delta_k) F_k \right. \\ & + (\bar{A}_k^{(\Sigma)} - A_k)^T (P_k + P'_k \delta_k) + (P_k + P'_k \delta_k) (\bar{A}_k^{(\Sigma)} - A_k) \\ & - \left( (\bar{B}_k - B_k) R_k^{-1} B_k^T P_k \right)^T (P_k + P'_k \delta_k) \\ & \left. - (P_k + P'_k \delta_k) \left( (\bar{B}_k - B_k) R_k^{-1} B_k^T P_k \right) \right] x_k + x_k^T \mathcal{O}(\delta^2) x_k \end{aligned} \quad (8.51)$$

where the higher-order terms from (8.48) have been absorbed into the  $\mathcal{O}(\delta^2)$  term (note – there are still no  $A$  or  $B$  terms with combined order greater than the order of  $\delta$  in this term). From (8.16), (8.17), (8.21), and (8.49) we have

$$\begin{aligned} \bar{A}_k^{(\Sigma)} &= \begin{bmatrix} A_{11}(t_k^{(1)}) & A_{12}(t_k^{(1)}) & \dots & A_{1n}(t_k^{(1)}) \\ A_{21}(t_k^{(2)}) & A_{22}(t_k^{(2)}) & \dots & A_{2n}(t_k^{(2)}) \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}(t_k^{(n)}) & A_{n2}(t_k^{(n)}) & \dots & A_{nn}(t_k^{(n)}) \end{bmatrix} \\ &= \begin{bmatrix} A_{11_k} + \delta_k^{(1)} A'_{11_k} + \dots & A_{12_k} + \delta_k^{(1)} A'_{12_k} + \dots & \dots & A_{1n_k} + \delta_k^{(1)} A'_{1n_k} + \dots \\ A_{21_k} + \delta_k^{(2)} A'_{21_k} + \dots & A_{22_k} + \delta_k^{(2)} A'_{22_k} + \dots & \dots & A_{2n_k} + \delta_k^{(2)} A'_{2n_k} + \dots \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1_k} + \delta_k^{(n)} A'_{n1_k} + \dots & A_{n2_k} + \delta_k^{(n)} A'_{n2_k} + \dots & \dots & A_{nn_k} + \delta_k^{(n)} A'_{nn_k} + \dots \end{bmatrix} \\ &= A_k + \Lambda_{\delta_k} A'_k + \mathcal{O}(\delta^2) \end{aligned} \quad (8.52)$$

where

$$\Lambda_{\delta_k} = \text{diag}\{\delta_k^{(1)}, \delta_k^{(2)}, \dots, \delta_k^{(n)}\} \quad (8.53)$$

so that

$$\bar{A}_k^{(\Sigma)} - A_k = \Lambda_{\delta_k} A_k' + \mathcal{O}(\delta^2) \quad (8.54)$$

Similarly, from the definition of  $\bar{B}_k$  and (8.50), we have

$$\bar{B}_k - B_k = \Lambda_{\delta_k} B_k' + \mathcal{O}(\delta^2) \quad (8.55)$$

Now (8.51) may be written

$$\begin{aligned} \Delta V_k \leq & s_{k+1} \delta_k \left\{ x_k^T \left[ \epsilon I + F_k^T P_k + P_k F_k + \delta_k (F_k^T P_k' + P_k' F_k) \right] x_k \right. \\ & + x_k^T \left[ (A_k')^T \Lambda_{\delta_k} P_k + P_k \Lambda_{\delta_k} A_k' \right] x_k \\ & - x_k^T \left[ \left( \Lambda_{\delta_k} B_k' R_k^{-1} B_k^T P_k \right)^T P_k + P_k \left( \Lambda_{\delta_k} B_k' R_k^{-1} B_k^T P_k \right) \right] x_k \left. \right\} \\ & + x_k^T \mathcal{O}(\delta^2) x_k \end{aligned} \quad (8.56)$$

Using the fact that, for symmetric  $M$ ,

$$x^T M x \leq \bar{\sigma}(M) x^T x \quad (8.57)$$

and

$$-x^T M x \leq -\underline{\sigma}(M) x^T x \quad (8.58)$$

where  $\bar{\sigma}$  and  $\underline{\sigma}$  denote the maximum and minimum singular values, respectively, and that

$$x^T \Lambda_{\delta_k} x \leq \delta_k x^T x \quad (8.59)$$

Equation (8.56) becomes

$$\begin{aligned} \Delta V_k \leq & s_{k+1} \delta_k x_k^T [\epsilon I + F_k^T P_k + P_k F_k] x_k + s_{k+1} \delta_k^2 x_k^T [F_k^T P_k' + P_k' F_k] x_k \\ & + s_{k+1} \delta_k^2 \bar{\sigma} \left[ (A_k')^T P_k + P_k A_k' \right] x_k^T x_k \\ & - s_{k+1} \delta_k^2 \underline{\sigma} \left[ \left( B_k' R_k^{-1} B_k^T P_k \right)^T P_k + P_k \left( B_k' R_k^{-1} B_k^T P_k \right) \right] x_k^T x_k \\ & + x_k^T \mathcal{O}(\delta^2) x_k \end{aligned} \quad (8.60)$$

$$= s_{k+1} \delta_k x_k^T [\epsilon I + F_k^T P_k + P_k F_k] x_k + x_k^T \mathcal{O}(\delta^2) x_k \quad (8.61)$$

where the last four terms in (8.60) have been combined into the  $\mathcal{O}$  term in (8.61), which is allowable since  $P'_k$ ,  $A'_k$ , and  $B'_k$  are guaranteed finite by our analyticity assumption. Now, it is easy to show that the SDRE (8.6) can be rearranged to give

$$F_k^T P_k + P_k F_k = -H_k^T H_k - P_k B_k R_k^{-1} B_k^T P_k \equiv -\bar{Q}_k \quad (8.62)$$

which is negative definite since  $H_k^T H_k > 0$  was assumed. Thus,  $\bar{Q}_k > 0$  and (8.61) becomes

$$\Delta V_k \leq s_{k+1} \delta_k x_k^T \left[ \epsilon I - \bar{Q}_k \right] x_k + x_k^T \mathcal{O}(\delta^2) x_k \quad (8.63)$$

and from (8.45) we are guaranteed that  $\epsilon I - \bar{Q}_k < 0$  for any  $k$ . Now, there exists a  $\delta_k$  sufficiently small so that the first term in (8.63) dominates and  $\Delta V_k < 0$ . Note that if  $\delta_k$  is chosen to ensure the right hand side of (8.63) is negative, then for any smaller  $\delta_k$  we still have  $\Delta V_k < 0$  since  $\bar{Q}_k$  remains constant and the possibly destabilizing contributions due to  $P'_k$ ,  $A'_k$ , and  $B'_k$  become smaller. Thus,  $\Delta V_k < 0$  for all  $t_r \in (t_k, t_{k+1}]$ , so that the intersample dynamics are well behaved.

Finally, we need to show there exist positive, real over- and underbounds for  $X_k$ , as defined by (8.43). From the assumption that  $\lim_{k \rightarrow \infty} P_k$  exists, we have that  $\|P_k\|_2$  is bounded above, while from the facts that  $P_k$  is positive definite and analytic, we are guaranteed  $\|P_k\|_2$  is bounded below away from 0. Now, for convergent  $P_k$ , (8.44) guarantees that  $s_k$  is the product of a finite (but possibly very large) number of positive real numbers for all  $k$ . Thus, the set of all  $s_k$  is a finite set of positive numbers which is guaranteed to possess maximum and minimum elements, both of which must be positive. The products of the lower and upper bounds on  $s_k$  and  $\|P_k\|_2$ , respectively, give  $\alpha$  and  $\beta$  in Theorem 8.2.1. ■

### 8.5 Discussion of Sampling Interval Size

With regard to practical application, (8.60) and (8.62) clearly show the interplay between the sample period size  $\delta_k$ , the state weighting matrix  $Q \equiv H^T H$ , the Riccati solution  $P$ , and the system dynamics parameters  $A$  and  $B$ . The interpretation is that the  $Q$  matrix needs to be selected ‘big enough’ and the sampling interval ‘small enough’ so that linear terms in the Taylor series expansions

of  $x_k$  dominate, and that additionally the rates of change of  $P$ ,  $A$ , and  $B$  are small enough to be dominated by the stabilizing effects of  $Q$ . It is interesting to note that this heuristic explanation of the theory agrees completely with the observations in [19] and the analytical results found for the case of nonlinear systems written as (8.2) in [59], in which conditions guaranteeing stability of quasi-linear parameter-varying systems are derived. It can be deduced from (8.4) that the rate of change of  $P$  will be constrained by the rates of change of  $A$ ,  $B$ ,  $R$ , and  $Q$ , so that the system dynamics are the determining factor if  $Q$  and  $R$  are chosen constant. For highly nonlinear systems (ones with high powers of  $x$  in the Taylor series expansions), we would expect large instantaneous time rates of change far from the origin due to the contribution from the partial with respect to  $x$ . For fixed  $Q$  and achievable sampling rate, by the above proof we would clearly expect such systems to have smaller domains of attraction than less highly nonlinear systems. Hence the claim of global stability for the nonlinear SDRE regulator would not be true for systems sampled at a finite rate, in general.

When actually selecting sampling intervals, several factors play a part. Recall from Theorem 8.3.1 that  $\delta_k$  must be selected small enough to make the state transition matrix representation valid, and we had (8.26) as a requirement for convergence of the infinite series representation. Additionally, from considering the stability proof in Theorem 8.4.1, we see that we need  $n\delta_k\bar{A}_k^2$  and  $n\delta_k\bar{A}_k\bar{B}_k$  to be very small compared to  $\bar{A}_k$  in order to ensure that the first-order terms dominate the state transition matrix expression. Also, we need  $\delta_k$  small enough to ensure that the Taylor series expansions of the system parameters are valid in a neighborhood of the sampling times, and moreover, that the constant and linear terms in these expansions dominate. Finally, the changes in  $P$ ,  $A$ , and  $B$  between sampling times must be very small compared to  $\bar{Q}$  at the given sampling time. Although several constraints exist, they all require  $\delta_k$  to be 'small', and thus are not disjoint but rather overlapping constraints which may be satisfied by choosing  $\delta_k$  small enough.

## 8.6 Conclusion

To summarize our results, we have shown that if  $\lim_{k \rightarrow \infty} P_k$  exists, and if the matrices  $A$ ,  $B$ ,  $R$  and  $H$  meet the following assumptions:

- analyticity in  $x$
- $\{A(x), B(x)\}$  stabilizable for all  $x$
- $H^T(x)H(x) > 0$  for all  $x$

then there exists a set of sampling times  $t_k$  such that the closed loop system using the sampled data state feedback SDRE regulator is asymptotically stable starting from any finite initial condition. From a physical point of view, stability will be enhanced by choosing larger state weighting matrices, at the expected price of higher control usage. For highly nonlinear systems, realistic constraints on sampling rates and control authority will severely limit domains of attraction.

# IX. Lyapunov Stability of Analytic Sampled Data Systems with Positive Semidefinite $Q$ Matrices

## 9.1 Introduction

In Chapter 8 it was shown that if the sampled data state feedback nonlinear regulation state-dependent Riccati equation (SDRE) yields a convergent solution, then, under some additional assumptions, the closed loop system will be semiglobally asymptotically stable. In this section we review the sampled data SDRE control algorithm and the assumptions used to prove the stability claim. In the remainder of the chapter we then show how the restrictive assumption of  $Q > 0$  for all  $x$  can be relaxed to  $Q \geq 0$  for all  $x$ , thereby greatly increasing both design flexibility and the set of systems for which stability may be proven.

We again consider regulation of continuous time, state feedback, input-affine, autonomous nonlinear dynamic systems of the form (8.1) with state vector  $x \in \mathcal{R}^n$ , control vector  $u \in \mathcal{R}^m$ , penalized variable  $z \in \mathcal{R}^s$ , and control penalty matrix  $\bar{R}(x)$  (assumed nonsingular for all  $x$ ). We assume that  $a$  and  $h$  are real-valued analytic functions of  $x$  on  $\mathcal{R}^n$ . We thus can again write (8.1) (nonuniquely) in state-dependent coefficient (SDC) form (8.2) where  $A$  and  $H$  are chosen to be analytic functions of  $x$ . The continuous control algorithm is again given by (8.3) and (8.4), and once more we assume that the pairs  $\{A(x), B(x)\}$  and  $\{H(x), A(x)\}$  are respectively stabilizable and detectable for all  $x$  to ensure global solutions of (8.4) and (8.6) exist. Upon substituting the control (8.3) into (8.1) we obtain the closed loop dynamics

$$\dot{x} = [A(x) - B(x)R^{-1}(x)B^T(x)P(x)]x \equiv F(x)x \quad (9.1)$$

where  $F(x)$  is guaranteed to be a Hurwitz matrix for all  $x$ . When actually implementing the above algorithm, we will use the sampled data control algorithm given by (8.5) and (8.6), which assumes perfect measurements of the state vector  $x$  and new control values at sampling times  $t_k$ . We thus again have a control which is held constant over the sampling interval  $[t_k, t_{k+1})$ .

## 9.2 Review of Stability Proof for $Q(x) > 0$

In Chapter 8 it is proven that under the discrete time control sequence (8.5), the closed loop dynamics admit a difference equation formulation

$$x_r = \Phi_{kr} x_k \quad \forall t_r \in [t_k, t_{k+1}] \quad (9.2)$$

provided the sampling intervals  $\delta_k = t_{k+1} - t_k$  are kept small enough, where  $\Phi_{kr}$  is an infinite matrix series which may be approximated by truncating higher-order terms. In order to prove global asymptotic stability of the closed loop system, it is assumed that  $P_k$  converges to some matrix limit  $\bar{P}$ , and  $H(x)$  is assumed to be nonsingular for all  $x$  so that  $Q_k = H_k^T H_k > \epsilon I > 0 \quad \forall k$ . This assumption on  $H$  strengthens the global detectability assumption on  $\{H(x), A(x)\}$  to global observability, so that the SDRE solution  $P_k > 0 \quad \forall k$  [76], and thus we have  $\bar{P} > 0$ . A positive definite Lyapunov function for the system is then defined for any  $t_r \in (t_k, t_{k+1}]$  as (recalling  $x(t_r) = x_r$ )

$$V(x_r) = s_r x_r^T P_r x_r \quad (9.3)$$

where  $P_r$  is the SDRE solution at  $x_r$ ,  $s_r$  is a positive scalar defined by the recursive expression

$$\begin{aligned} s_r &= \frac{x_k^T P_k x_k}{x_k^T P_r x_k} s_k \quad (if \Delta_r \geq \epsilon) \\ &= s_k \quad (otherwise) \end{aligned} \quad (9.4)$$

with  $s_0 = 1$  and where  $\epsilon$  is an arbitrarily small positive tolerance selected so that

$$\epsilon < \min_k [\underline{\lambda}(H_k^T H_k)] \quad (9.5)$$

where  $\underline{\lambda}$  denotes the minimum eigenvalue, and where

$$\begin{aligned} \Delta_r &= \frac{x_k^T (P_r - P_k) x_k}{\delta_r x_k^T x_k} \quad (if \ x_k \neq 0) \\ &= 0 \quad (if \ x_k = 0) \end{aligned} \quad (9.6)$$

for  $\delta_r = t_r - t_k$ . It is then shown that the Lyapunov function variation over the interval from  $t_k$  to  $t_r$ ,

$$\Delta V_{kr} \equiv V_r - V_k$$

$$\begin{aligned}
&= x_k^T (s_r P_r - s_k P_k) x_k + s_r \delta_r x_k^T [F_k^T P_k + P_k F_k] x_k + x_k^T \mathcal{O}(\delta_k^2) x_k \\
&\leq -s_r \delta_r x_k^T [-\epsilon I + Q_k + P_k B_k R_k^{-1} B_k^T P_k] x_k + x_k^T \mathcal{O}(\delta_k^2) x_k
\end{aligned} \tag{9.7}$$

is guaranteed negative for all  $t_r \in (t_k, t_{k+1}]$ , and for all  $k$ , provided the system matrices are analytic in  $x$  and  $\delta_k$  is selected small enough. Note that the final term in (9.7) stands for higher-order terms in  $\delta_k$ . Furthermore, convergence of  $P_k$  guarantees boundedness of  $P_k$  and  $s_k$  from above and below (away from zero), so that (9.3) is also radially unbounded, thus allowing the claim of global asymptotic stability by standard Lyapunov type arguments.

### 9.3 Stability for $Q(x) \geq 0$

#### 9.3.1 LaSalle's Invariance Principle and Redefined Lyapunov Function

The global nonsingularity assumption on  $H(x)$  is restrictive, and we thus seek means of relaxing it by allowing  $\Delta V_{kr} \leq 0$  and invoking LaSalle's Invariance Principle [44, 69] to maintain the stability claim. For ease of reference we thus present a discrete time version of this theorem, taken from [20]. A prerequisite is the definition of an *invariant set*. A set  $M \subseteq \mathcal{R}^n$  is said to be invariant with respect to (9.2) if, for some  $t_0 \geq 0$ ,

$$x(t_0) = x_0 \in M \Rightarrow x(t_k) = x_k \in M \quad \forall t_k \in \mathcal{R}_+ \tag{9.8}$$

where  $x_k$  represents the solution of (9.2) at time  $t_k$  starting from  $x_0$  at time  $t_0$ . We now give the theorem.

**Theorem 9.3.1 (Discrete Lasalle Invariance)** *Let  $\Omega$  be an invariant set of (9.2), and let  $V : \Omega \rightarrow \mathcal{R}_+$  be a continuous function  $V(x)$  such that  $\Delta V_k = V_{k+1} - V_k \leq 0 \quad \forall x_k \in \Omega$ . Also, let  $E = \{x_k \in \Omega \mid \Delta V_k = 0\}$ , and let  $M$  be the maximal invariant set contained in  $E$ , i.e., the union of invariant sets contained in  $E$ . Then every bounded solution  $x(t_k)$  starting in  $\Omega$  converges to  $M$  as  $t_k \rightarrow \infty$ .*

**Proof:** See [20]. ■

It is well known [11, 69] that Theorem 9.3.1 offers increased flexibility in proving stability of autonomous systems, because it does not require  $V$  to be strictly positive definite, nor does it require  $\Delta V$  to be strictly negative, and it also relaxes the  $C^1$  assumption on  $V$  to continuity only. This freedom comes at the expense of somehow establishing boundedness of system trajectories, however. In our case we have boundedness by virtue of the fact that, since  $P$  is a continuous (analytic) function of  $x$ , convergence of  $P_k$  implies convergence of  $x_k$ , which implies boundedness of  $x_k$  [1]. Also note that under our assumption of sufficiently small  $\delta_r$ , the analyticity of  $P$  may be used to write  $P_r \approx P_k + \delta_r P'_k$  where  $P'_k$  is the matrix of time derivatives of  $P$  evaluated at  $t_k$ , so that for  $x_k \neq 0$  (9.6) can be replaced by

$$\Delta_r = \frac{x_k^T P'_k x_k}{x_k^T x_k} \quad \forall t_r \in [t_k, t_{k+1}) \quad (9.9)$$

Thus, for sufficiently fast sampling,  $\Delta_r$  is the same for all values of time in the sampling interval, so that only one comparison versus  $\epsilon$  need be made per sampling interval, and thus either  $s_r = s_k$  for all  $t_r \in [t_k, t_{k+1})$  or  $s_r$  is determined from the first line of (9.4) for all  $t_r \in [t_k, t_{k+1})$ . In the first case,  $s$  is constant and thus continuous over the sampling interval, while in the second case the continuity of  $P$  gives  $s$  as a continuous function of  $x$ . Thus,  $V$  is continuous as required for use in Theorem 9.3.1. However, with positive semidefinite  $Q = H^T H$ , we cannot guarantee satisfaction of (9.5), so that the definition of the Lyapunov function must be altered to guarantee negativity of its variation, or establish stability via Lasalle's Invariance Principle. We therefore maintain (9.3), but now we define  $s_r$  for  $t_r \in (t_k, t_{k+1}]$  according to

$$s_r = \frac{x_k^T P_k x_k}{x_k^T P_r x_k} s_k \quad (\text{if } x_k^T P'_k x_k \geq x_k^T [Q_k + P_k B_k R_k^{-1} B_k^T P_k] x_k) \quad (9.10)$$

$$= s_k \quad (\text{otherwise}) \quad (9.11)$$

where  $P'_k$  is as above and again we take  $s_0 = 1$ . By arguments analagous to the above, such a definition gives a continuous  $V(x)$ , and has the result of changing the Lyapunov function variation to

$$\Delta V_{kr} = -s_r \delta_r x_k^T [Q_k + P_k B_k R_k^{-1} B_k^T P_k] x_k + x_k^T \mathcal{O}(\delta_k^2) x_k \quad (9.12)$$

for (9.10) so that the  $\epsilon I$  term in (9.7) is no longer present. For (9.11) we get a resulting  $\Delta V_{kr}$  as given in (9.12), plus the possible addition of an uncanceled nonzero  $s_r x_k^T (P_r - P_k)x_k$  term, whose contribution to  $\Delta V_{kr}$  is not large enough to ruin its negativity. In simple terms, we leave the  $x^T P x$  term alone when the first-order terms produce a negative  $\Delta V$ , and we introduce a scaling to eliminate the growth of  $x^T P x$  when it is sufficiently large to be detrimental to stability.

### 9.3.2 Characterization of $E$

We thus now have a continuous Lyapunov function with nonpositive variation so that we are motivated to study invariant sets of (9.2) such that for some  $k'$  the necessary condition

$$s_r x_{k'}^T [Q_{k'} + P_{k'} B_{k'} R_{k'}^{-1} B_{k'}^T P_{k'}] x_{k'} = 0 \quad (9.13)$$

holds. To incorporate the notion of invariance and strengthen (9.13) to a necessary and sufficient condition, we need to identify when  $\Delta V_{kr} = 0 \forall t_r \in (t_k, t_{k+1}]$  and for all  $k \geq k'$ . Thus, recall that

$$\begin{aligned} \Delta V_{kr} &= V_r - V_k \\ &= s_r x_r^T P_r x_r - s_k x_k^T P_k x_k \end{aligned} \quad (9.14)$$

$$= x_k^T [s_r \Phi_{kr}^T P_r \Phi_{kr} - s_k P_k] x_k \quad (9.15)$$

where  $\Phi_{kr}$  is the state transition matrix from  $x_k$  to  $x_r$  as defined in (9.2), so that from (9.14) we have the four possibilities

- i.  $\Delta V_{kr} = 0$  if  $x_k \in \mathcal{N}\{P_k\}$  and  $x_r \in \mathcal{N}\{P_r\}$  where  $\mathcal{N}$  represents the null space of a matrix, while from (9.15) and (9.11) we have
- ii.  $\Delta V_{kr} = 0$  if  $\Phi_{kr} = I$
- iii.  $\Delta V_{kr} = 0$  if  $x_r \neq x_k$  and  $s_r x_r^T P_r x_r = s_k x_k^T P_k x_k = c \neq 0$ , and
- iv.  $\Delta V_{kr} = 0$  if  $s_k = 0$

at least one of which must be satisfied for all  $t_r \in [t_k, t_{k+1}]$  and all  $k \geq k'$  for invariance to hold.

Thus, items i, ii, iii, and iv above give a complete characterization of the set  $E$  in Theorem 9.3.1.

Now, considering item ii, from (9.1) we have for sufficiently small  $\delta_r$  that  $\Phi_{kr} = I$  iff  $F(x_r)x_r = 0 \forall t_r \in [t_k, t_{k+1}]$ , while item i above implies that  $P_r x_r = x_r^T P_r = 0$  for all  $t_r$  in the interval. Item iii requires for all  $t_r \in [t_k, t_{k+1}]$  that  $F_r x_r \neq 0$ ,  $P_r x_r \neq 0$ , and the trajectory evolves in such a way that  $s_r x_r^T P_r x_r$  remains constant and nonzero. For sufficiently small  $\delta_r$ , it is easily shown that this nonzero rate of change condition on the quadratic form is equivalent to orthogonality of the two nonzero vectors  $Fx$  and  $Px$ . Thus, by considering the set

$$E = \{x_k \in \mathcal{R}^n \mid x_k^T P_k F_k x_k = 0 \text{ or } s_k = 0\} \quad (9.16)$$

we have an alternate but completely equivalent characterization of  $E$ . Now, since  $x^T P F x = x^T F^T P x$  we may consider

$$E = \{x_k \in \mathcal{R}^n \mid x_k^T [F_k^T P_k + P_k F_k] x_k = 0 \text{ or } s_k = 0\} \quad (9.17)$$

the first part of which is well known to be equivalent to the set of all  $x_k$  such that (9.13) holds (with the appropriate change in subscripts), by simple manipulation of the SDRE (8.4). Thus, we seek to identify all invariant sets of (9.1) such that

$$E = \{x \in \mathcal{R}^n \mid x^T [Q + PBR^{-1}B^T P]x = 0 \text{ or } s(x) = 0\} \quad (9.18)$$

and guarantee asymptotic stability to the origin of all trajectories contained in these invariant sets.

### 9.3.3 Nonlinear Observability and Invariant Sets

Let us consider the first expression in (9.18), namely

$$x^T [Q + PBR^{-1}B^T P]x = 0 \quad (9.19)$$

which is clearly seen to be the sum of two positive semidefinite terms, recalling that  $Q = H^T H$ . Thus, for (9.19) to be zero, we must have both terms equal to zero. By also noting that  $u = -R^{-1}B^T Px$ , (9.19) can be rearranged to give

$$(Hx)^T (Hx) + u^T R u = 0 \quad (9.20)$$

and recalling that  $R > 0$ , we thus must have  $x \in \mathcal{N}[H]$  and  $u = 0$  for (9.20) to hold, so that what we really desire to identify are the invariant sets of the continuous time open loop dynamics contained in the nullspace of  $H$ , i.e., all  $x$  such that  $x \in \mathcal{N}[H]$  and  $x$  is invariant with respect to

$$\dot{x} = a(x) = A(x)x \quad (9.21)$$

Such sets may occur in two ways: as the result of subsets of observable states, or as the result of subsets of unobservable states. By unobservable states we mean invariant manifolds in the state space on which  $y$  does not change, even though the states  $x$  are changing. Such invariant manifolds admit a coordinate transformation so that the output does not depend on the unobservable states, where the number of unobservable states is equal to the dimension of the manifold [26]. Thus, nontrivial invariant unobservable manifolds will have dimension one or greater, so that manifolds in the observable space on which the output remains constant will consist of only single isolated points (zero dimensional manifolds). We talk about subsets of the above manifolds because even if the output remains constant on such a manifold, we still are interested in sets where  $h$  remains constant and equal to zero. The result of this is that, for a completely observable system, even if we regulate the output to zero, we still must ensure that isolated invariant points such that  $h = 0$  do not exist, or else we cannot guarantee that all the states are driven to zero. For a system with a nontrivial unobservable space, even if all observable states are driven to zero by regulation of the output, the remaining unobservable states do not affect the output and so are not compelled to go to zero, and thus closed loop stability may be lost. We must therefore identify the unobservable states and ensure all trajectories constrained to the unobservable manifold converge to zero. It is known how to characterize unobservable manifolds, and, in fact, in [33], a recursive algorithm for this is given. We repeat it here, referring the reader to [33] for additional details. Let  $\tau_i(x)$  be covector fields spanning the row space of  $dh(x) = \frac{\partial h}{\partial x}$  for all  $x$ , and let  $\Omega_0 = \text{span}[\tau_i]$ . Then let

$$\Omega_p = \Omega_{p-1} + L_a \Omega_{p-1} \quad (9.22)$$

where  $L_a \Omega_{p-1}$  represents the Lie derivative of each covector field in  $\Omega_{p-1}$  with respect to  $a$ , and the  $+$  sign represents the subspace sum, i.e., the sum of the spans. If the above procedure converges so

that  $\Omega_p = \Omega_{p-1}$  for some  $p$ , then the unobservable states of (9.21) through  $h$  are a subset of

$$\mathcal{O}_{nl}^\perp = \{x \in \mathcal{R}^n \mid \Omega_p x = 0\} \quad (9.23)$$

which is the nullspace of a matrix function of  $x$  (a distribution), so that these sets may vary from point to point in the state space. However, if the dimension  $d$  of  $\Omega_p$  is constant for all  $x$ , i.e.,  $\Omega_p$  is a nonsingular distribution, then  $\mathcal{O}_{nl}^\perp$  will be a smooth manifold of dimension  $n - d$ , and the above mentioned transformation to invariant observable and unobservable states may be performed. It should be stressed that nonsingularity of  $\Omega_p$  is critical to construction of the invariant, observable and unobservable manifolds. If  $\Omega_p$  is singular at some points in the state space, the decomposition might still be performable, but now requires additional considerations such as using  $\text{smt}(\Omega_p)$ , the largest smooth distribution contained in  $\Omega_p$ , in place of  $\Omega_p$  itself [33]. In either case it is known that  $\Omega_{n-1}$  gives a characterization of the observable space  $\mathcal{O}_{nl}$  on an open and dense subset of  $\mathcal{R}^n$ , so that often  $p = n$  in (9.23). By assuming this and applying some properties of Lie derivatives, we can express (9.23) alternatively as

$$\mathcal{O}_{nl}^\perp = \mathcal{N} \begin{bmatrix} dHx \\ dL_a^1 Hx \\ \vdots \\ dL_a^{n-1} Hx \end{bmatrix} \quad (9.24)$$

where  $L_a^i$  is a shorthand notation for differentiation  $i$  times along the vector field  $a$ . We thus see that a sufficient condition for there not to exist any invariant sets of (9.21) unobservable through  $h$  other than  $x = 0$  is for the rank of (9.24) to be equal to  $n$ , the dimension of the state space, for all  $x$ . This leads us to our first sufficient condition for asymptotic closed loop stability with  $Q \geq 0$  as opposed to  $Q > 0$ , as stated below.

**Theorem 9.3.2** *Consider (8.1) written as (8.2), with the system matrices in (8.2) globally stabilizable/detectable and analytic with respect to  $x$ . Assume the sampled data SDRE control law (8.6), (8.5) is applied to (8.2) where  $\delta_k$  is selected appropriately small to guarantee dominance of first-order*

terms in the expression for  $\Delta V_k$ , assume  $\lim_{k \rightarrow \infty} s_k \neq 0$ , and further assume  $\text{rank}[\Omega_p] = n$  for some  $p \geq 0$  and for all  $x$ . Then the closed loop system is asymptotically stable.

**Proof:** The proof follows directly from Theorem 9.3.1, noting that the assumptions guarantee that the set  $M$  consists only of the zero vector. The assumption on  $s_k$  takes the given form because  $s_k$  is guaranteed positive for any finite  $t_k$ , and thus it cannot cause  $\Delta V_k$  to equal zero except as  $k$  tends to infinity. Now, recall that it was mentioned that invariant sets in the nullspace of  $H$  might also occur on subsets of the observable space. However, by definition such sets must consist of isolated points, since otherwise they would belong to locally invariant unobservable manifolds, and thus would not be in the observable space. For isolated points to be invariant and in the nullspace of  $H$  we must have at such points that  $h = Hx = 0$  and  $\dot{h} = [H + \frac{\partial H}{\partial x}]Ax = 0$ , as well as having higher-order derivatives equal to zero. However, it may be shown that for analytic systems such cannot be the case if  $\text{rank}[\Omega_p] = n$ . This is a consequence of an analytic function being zero at  $x = 0$  and having nonzero, continuous gradient everywhere else. Thus, Theorem 9.3.2 precludes the need to consider isolated invariant points in the nullspace of  $H$ , so that the theorem does indeed provide a sufficient condition for asymptotic closed loop stability. ■

At this point we observe that there are two sources of possible conservatism in the above theorem. Note first of all that if  $h$  is a linear function of  $x$ , then the first entry in  $\Omega_p$  is  $H$ , so that any  $x$  in  $\Omega_p^\perp$  necessarily also has  $h = Hx = 0$ , and all the unobservable states are also in the nullspace of  $H$ . However, if  $h$  is not a purely linear function of  $x$ , then the first entry of  $\Omega_p$  is  $dHx = H + \frac{\partial H}{\partial x}$  so that  $x \in \Omega_p^\perp$  does not necessarily imply that  $x$  is in the nullspace of  $H$ . The result of this is that for nonlinear  $h$  the invariant unobservable sets in the nullspace of  $H$  truly are, in general, a subset of the overall invariant unobservable manifold, and thus the unobservable manifold may contain points not in the set  $E$ , which is what we are trying to identify. However, restricting a set to contain only the zero vector as Theorem 9.3.2 does also necessarily restricts any of its subsets to equal the zero vector. The second source of possible conservatism is that the above analysis assumes that  $u = 0$ ,

and proceeds to identify invariant sets based on that assumption. It is therefore of interest to try to identify when this assumption is satisfied; i.e., for what  $x$  is  $u(x) = 0$ ?

### 9.3.4 Factored Observability and Control Action

The analysis is aided by decomposition of (8.2) into linearly as opposed to nonlinearly observable and unobservable parts, where linear here means in terms of the pointwise linearization matrices  $A(x)$  and  $H(x)$ . It is well known from linear systems theory [41], that the LTI system

$$\begin{aligned}\dot{x} &= Ax + Bu \\ z &= Hx\end{aligned}\tag{9.25}$$

has an invariant, unobservable space given by

$$\mathcal{O}_l^\perp = \mathcal{N} \begin{bmatrix} H \\ HA \\ \vdots \\ HA^{n-1} \end{bmatrix}\tag{9.26}$$

and that for  $\{A, B\}$  stabilizable,  $x \in \mathcal{O}_l^\perp$  iff  $x \in \mathcal{N}[P]$  [50], where  $P$  is the positive semidefinite stabilizing solution to the algebraic Riccati equation

$$A^T P + PA - PBR^{-1}B^T P + Q = 0\tag{9.27}$$

Thus,  $Px = 0$  for every  $x \in \mathcal{O}_l^\perp$  so that  $u(x) = -R^{-1}B^T Px = 0$  for all  $x \in \mathcal{O}_l^\perp$ . It is also true [33] that for an LTI system, (9.24) simplifies to (9.26), so that  $\mathcal{O}_l^\perp$  is an invariant set of (9.25) in which  $u = 0$  and  $Hx = 0$ . Thus, for an LTI system,  $\mathcal{O}_l^\perp$  as defined in (9.26) completely defines the unobservable subspace, which by Lasalle's Invariance principle, requires additional assumptions for stability. In LTI linear quadratic regulator theory this is handled by requiring the invariant unobservable dynamics to be stable (the detectability assumption). If we now generalize (9.26) to

state-dependent factorizations, we find that if  $x$  belongs to the set

$$\mathcal{O}_f^\perp = \mathcal{N} \begin{bmatrix} H(x) \\ H(x)A(x) \\ \vdots \\ H(x)A^{n-1}(x) \end{bmatrix} \quad (9.28)$$

then we still have  $u(x) = 0$  and  $H(x) = 0$ , but we are not guaranteed that  $\mathcal{O}_f^\perp$  is an invariant set of (9.21) because of possible discrepancies between  $\mathcal{O}_f^\perp$  and  $\mathcal{O}_{nl}^\perp$ . These discrepancies arise as a result of a general nonequivalence between the  $H(x)A^i(x)$  and of the Lie derivatives  $dL_a^i Hx$ , and hence of their nullspaces (see Chapter 6 for proofs of nonequivalence for the dual case of controllable sets). Note that the above analysis again points to the necessity of precluding isolated invariant points in the nullspace of  $H$ , for if  $x$  is such that  $A(x)x = 0$  and  $H(x)x = 0$ , i.e.,  $x$  is an open loop equilibrium in the nullspace of  $H$ , then clearly,  $x \in \mathcal{O}_f^\perp$  and  $x \in M$ . As an immediate consequence of these arguments we obtain the following theorem regarding a necessary condition for global asymptotic stability of SDRE nonlinear regulators with positive semidefinite state weighting matrices.

**Theorem 9.3.3** *Consider (8.1) written as (8.2). Suppose multiple open loop equilibria exist such that  $a(x) = A(x)x = 0$  for  $x \neq 0$ , and define  $A_0$  to be the set of all such  $x \neq 0$ . Then, in order for the closed loop system (9.1) to be globally asymptotically stable, we must have  $Hx \neq 0 \forall x \in A_0$ .*

**Proof:** The proof follows from observing that  $x \in A_0$  and  $Hx = 0$  implies  $x \in \mathcal{O}_f^\perp$ , so that  $u(x) = 0$ . Thus, we have  $F(x)x = A(x)x + B(x)u(x) = 0$  so that  $x \in M$  and in fact  $x$  is a closed loop equilibrium point other than  $x = 0$ , contradicting global asymptotic stability of the origin. ■

We observe that choosing a globally nonsingular  $H(x)$  obviously satisfies the requirements of Theorem 9.3.3, preventing any difficulties from multiple open loop equilibria. For singular  $H(x)$ , we have the following sufficient condition for satisfaction of the requirements of Theorem 9.3.3.

**Theorem 9.3.4** *Consider (8.1) written as (8.2), and assume  $\{H(x), A(x)\}$  is detectable for all  $x$ . Then the necessary condition for global asymptotic stability given in Theorem 9.3.3 is satisfied.*

**Proof:** Recall that global detectability of  $\{H(x), A(x)\}$  implies that for all  $\lambda, x$  and  $z$  such that  $A(x)z = \lambda z$  and  $\operatorname{Re}(\lambda) \geq 0$ ,  $H(x)z \neq 0$ . Let  $\lambda = 0$ . Then  $A(x)z = 0$  implies  $H(x)z \neq 0$ , so that the theorem is proven. ■

### 9.3.5 Characterization of $M$

To this point we have seen that if  $x \in \mathcal{O}_f^\perp$ , then  $u(x) = 0$ . However, we do not necessarily have the reverse implication, which means that we do not know the set of all  $x$  such that  $u(x) = 0$ , and thus we do not know which  $x \in \mathcal{O}_{nl}^\perp$  are actually troublesome from a stability viewpoint. However, if we assume that the system is nonlinearly stabilizable and  $s_k$  remains nonzero, we do not have to be concerned with  $x$  outside of  $\mathcal{O}_f^\perp$ , as proven in the following theorem.

**Theorem 9.3.5** *Consider (8.1) written as (8.2), assume  $\lim_{k \rightarrow \infty} s_k \neq 0$ , and assume that (8.1) is nonlinearly stabilizable, as defined in Section 4.7. Then the set  $M$  in Theorem 9.3.1 for the SDRE regulator, i.e., the union of invariant sets of (9.21) contained in the nullspace of  $H$ , is contained in the intersection of  $\mathcal{O}_f^\perp$  and  $\mathcal{O}_{nl}^\perp$ , which we call  $I_{\mathcal{O}^\perp}$ .*

**Proof:** We prove only the case of singular  $H(x)$  for some  $x \neq 0$ , since the case for globally nonsingular  $H(x)$  is trivial, the intersection being the zero vector. Obviously, if  $x$  is an element of both  $\mathcal{O}_f^\perp$  and  $\mathcal{O}_{nl}^\perp$ , then  $Hx = 0$  and  $u(x) = 0$  so that  $x \in E$  and, in fact,  $x$  remains in the nullspace of  $H$  under (9.21), so that  $x \in M$ . Now, we have previously shown that  $M \subseteq \mathcal{O}_{nl}^\perp$ . From this it is clear that we cannot have  $x \in \mathcal{O}_f^\perp$  and  $x \in M$ , without having  $x \in \mathcal{O}_{nl}^\perp$  as well. We therefore need to prove only that  $x$  cannot be an element of both  $M$  and  $\mathcal{O}_{nl}^\perp$  without being an element of  $\mathcal{O}_f^\perp$ , assuming (8.1) is nonlinearly stabilizable. Thus, assume  $x \in \mathcal{O}_{nl}^\perp$  and  $x \in M$ . This implies  $Hx = 0$ ,  $u(x) = 0$  and  $x$  remains in the nullspace of  $H$  under (9.21). Now, suppose  $x$  is not in  $\mathcal{O}_f^\perp$ . Note that this excludes any  $x \in A_0$ , where  $A_0$  is defined as in Theorem 9.3.3, since any  $x \in A_0$  is automatically in both  $\mathcal{O}_{nl}^\perp$  and  $\mathcal{O}_f^\perp$  if  $Hx = 0$  as well. Thus,  $A(x)x \neq 0$  so that  $x$  may not be an isolated closed loop equilibrium point, and also we have  $P(x)x \neq 0$ , but  $u(x) = 0$ . Now, recalling  $u = -R^{-1}B^TPx$ , we see that  $u = 0$  iff  $B^TPx = 0$ , since  $R$  is globally nonsingular. From the previous discussion we

have  $x \neq 0$  and  $Px \neq 0$ , so that  $Px \neq 0$  must belong to an invariant set of (9.21) in the nullspace of  $B^T$ . Since, alternatively, this is an invariant set of (9.21) such that  $x^T PB = 0$ , we see that this may actually be considered a nonlinear controllability question. If (8.1) is weakly controllable on  $\mathcal{R}^n$ , then there is no nonzero vector  $w^T = x^T P$  which remains in the left nullspace of  $B$  for all  $x$  under (9.21). Thus, global weak controllability of (8.1) may be assumed in place of the weaker assumption of nonlinear stabilizability in the theorem statement. Nonlinear stabilizability of (8.1), however, is sufficient to guarantee that trajectories contained in the left nullspace of  $B$  converge to zero. Hence  $x^T P$  converges to zero, which is a contradiction, so that no such  $x$  outside of  $\mathcal{O}_f^\perp$  may be in  $M$ , and the theorem is proven. ■

This leads us to additional sufficient conditions for asymptotic closed loop stability with  $Q \geq 0$ , as given in the following theorems.

**Theorem 9.3.6** *Consider (8.1) written as (8.2), with (8.1) nonlinearly stabilizable and with the system matrices in (8.2) globally stabilizable/detectable and analytic with respect to  $x$ . Assume the sampled data SDRE control law (8.6), (8.5) is applied to (8.2) where  $\delta_k$  is selected appropriately small to guarantee dominance of first-order terms in the expression for  $\Delta V_k$ , assume  $\lim_{k \rightarrow \infty} s_k \neq 0$ , and further assume  $\mathcal{O}_f^\perp = 0$ . Then the closed loop system is asymptotically stable.*

**Proof:** The proof follows from Theorem 9.3.1 analogously to that of Theorem 9.3.2, where we have ensured  $M = 0$  by the assumptions of nonlinear stabilizability and  $\mathcal{O}_f^\perp = 0$  and invoking Theorem 9.3.5. ■

**Theorem 9.3.7** *Consider (8.1) written as (8.2), with (8.1) nonlinearly stabilizable and with the system matrices in (8.2) globally stabilizable/detectable and analytic with respect to  $x$ . Assume the sampled data SDRE control law (8.6), (8.5) is applied to (8.2) where  $\delta_k$  is selected appropriately small to guarantee dominance of first-order terms in the expression for  $\Delta V_k$ . Assume  $\lim_{k \rightarrow \infty} s_k \neq 0$ , and further assume  $I_{\mathcal{O}^\perp} = 0$ . Then the closed loop system is asymptotically stable.*

**Proof:** The proof follows from Theorem 9.3.1 analogously to that of Theorem 9.3.6, where we have ensured  $M = 0$  by the assumptions of nonlinear stabilizability and  $I_{O^\perp} = 0$  and invoking Theorem 9.3.5. ■

**Theorem 9.3.8** *Consider (8.1) written as (8.2), with (8.1) nonlinearly stabilizable and with the system matrices in (8.2) globally stabilizable/detectable and analytic with respect to  $x$ . Assume the sampled data SDRE control law (8.6), (8.5) is applied to (8.2) where  $\delta_k$  is selected appropriately small to guarantee dominance of first-order terms in the expression for  $\Delta V_k$ , assume  $\lim_{k \rightarrow \infty} s_k \neq 0$ , and further assume that for  $x$  restricted to  $I_{O^\perp}$ , (9.21) is asymptotically stable. Then the closed loop system is asymptotically stable.*

**Proof:** The proof follows from Theorem 9.3.1 analogously to that of Theorem 9.3.7, where the assumptions ensure that trajectories confined to  $M$  are asymptotically stable. ■

### 9.3.6 The Effect of Factored Controllability

The above analysis shows that, as long as the system considered is nonlinearly stabilizable, potentially troublesome sets from a stability perspective arise from the invariant open loop unobservable (unpenalized) states. Nonlinear stabilizability also allows this set of concern to be reduced by excluding  $x$ 's not also in the factored unobservable space. From these statements we see that nonlinear controllability (alternatively, stabilizability) plays an important role in our stability analysis. To conclude this discussion, let us make a comment regarding factored uncontrollable sets of (8.2), and their impact on closed loop stability analysis. Analogous to the above derivation of observable and unobservable sets, it is possible to obtain characterizations of factored controllable and uncontrollable sets for (8.2) (see Chapter 6). In fact, it can be shown [41] that based on the factored uncontrollable set defined by

$$\mathcal{C}_f^\perp \equiv \{x \in \mathcal{R}^n \mid x^T M_{cf}(x) = x^T [B(x) A(x)B(x) \cdots A(x)^{n-1} B(x)] = 0\} \quad (9.29)$$

(8.2) may be decomposed at each  $x$  into the form

$$\begin{aligned}\dot{x}_1 &= A_{11}(x)x_1 + A_{12}(x)x_2 + B_1(x)u \\ \dot{x}_2 &= A_{22}(x)x_2\end{aligned}\tag{9.30}$$

where  $x_1$  is in the factored controllable subspace  $\mathcal{C}_f$ ,  $x_2$  is in the factored uncontrollable subspace  $\mathcal{C}_f^\perp$ , and  $\{A_{11}(x), B_1(x)\}$  is a controllable pair. Thus, if  $x \in \mathcal{C}_f^\perp$ , then  $x^T B(x) = 0$  so that  $x$  has a component unaffected by the control ( $x_2$ ), and this component could be troublesome from a stability point of view if the above transformation yields an invariant manifold within the pointwise  $x_2$  space, and the appropriate open loop dynamics confined to this manifold are not stable. This eventuality is precluded, however, by the assumption of nonlinear stabilizability. The impact of a globally stabilizable parametrization on these issues is that, analogous to Theorem 9.3.4, using such a factorization prevents isolated uncontrollable points from being invariant. This is because such a globally stabilizable parametrization guarantees that any  $x$  such that  $x^T B = 0$  is such that  $Ax$  is nonzero, so that such an  $x$  does not belong to the set  $E$ , and hence not to  $M$  either, unless it belongs to a larger, invariant, uncontrollable set. Note that uncontrollable points as determined from (9.29), (9.30) which are also in  $\mathcal{O}_{nl}^\perp$  add nothing new to the analysis, since the effect of  $\mathcal{O}_{nl}^\perp$  has already been taken into account. From the above considerations, we see that uncontrollable states are important from a stability analysis viewpoint, but, given a globally stabilizable factorization, they are determined from the true nonlinear controllability, and not from the factored system. Their impact is in determining whether  $P_k$  converges for  $x \in \mathcal{O}_f$ , identifying what states must be placed in the nullspace of  $H$  due to lack of controllability/stabilizability, and whether any such uncontrollable dynamics in  $M$  are stable. To summarize, we have, assuming  $\{A(x), B(x)\}$  stabilizable for all  $x$ , that

- if  $x \in \mathcal{C}_f^\perp$  and  $x \in \mathcal{O}_{nl}^\perp$ , then  $x$  may be in  $M$  as determined by Theorem 9.3.5. Overall stability requires stability of open loop trajectories contained in  $M$  (zero state detectability);
- if  $x \in \mathcal{C}_f^\perp$ , and  $x$  is not in  $\mathcal{O}_{nl}^\perp$ , then  $x$  is not in  $M$ ;

- if the system is nonlinearly stabilizable, then  $\mathcal{O}_{nl}^\perp$  in the above may be replaced with  $\mathcal{O}_f^\perp$  or  $I_{\mathcal{O}^\perp}$ . In this case, if  $x$  is in the true invariant nonlinearly uncontrollable subspace, then similar to the second item above,  $x$  must be in  $\mathcal{O}_f^\perp$  for global closed loop stability to hold.

### 9.3.7 Convergence of $P_k$ and $s_k \neq 0$

Finally, let us consider the condition  $s(x) = 0$  of Section 9.3.2. By definition,  $s(x_r)$  is either equal to  $s_k$ , the value of  $s$  at the previous sampling time, or the ratio of two positive numbers times  $s_k$ . Thus, as long as  $t_k$  is finite, there are a finite number of positive terms in the recursion for  $s_k$ , so that  $s(x_r)$  is finite, bounded, and positive. The same is true if  $x_k$  converges to zero. By (9.11), if  $x$  becomes small enough (enters the domain of attraction of the linearized closed loop system, which is guaranteed asymptotically stable near the origin), then  $P$  becomes and remains essentially constant so that (9.10) will no longer hold. At some point in this case  $s_r = s_k$  for all  $t_r$  from that time on, where  $s_k$  has been computed as a product of a finite number of positive terms. We therefore are concerned only with preventing  $s(x_k)$  from tending to zero as  $t_k \rightarrow \infty$  for positive definite  $P_k$  and  $x_k$  not tending to zero, as stated in the above theorems. We restrict ourselves to positive definite  $P_k$  for sufficiently large  $k$  because if  $P_k$  is only positive semidefinite and we have convergence to an  $\bar{x} \neq 0$  in the nullspace of  $P_k$ , then we already know that the Lyapunov variation  $\Delta V_k$  will equal zero, and we do not care if  $s_k$  tends to zero or not. Thus, we now show that  $s_k$  will not tend to zero under the assumptions of global analyticity and stabilizability/detectability of the system matrices and convergence of  $P_k$  to  $\bar{P} > 0$ , with assumed constant sampling interval.

**Theorem 9.3.9** *Consider (8.1) written as (8.2), with the system matrices in (8.2) globally stabilizable/detectable and analytic with respect to  $x$ . Assume that the sampled data SDRE control law (8.6), (8.5) is applied to (8.2) where  $\delta_k = \delta$  is selected appropriately small to guarantee dominance of first-order terms in the expression for  $\Delta V_k$ , and assume that the closed loop trajectory converges to a point in  $\mathcal{O}_f$ , so that  $\lim_{k \rightarrow \infty} P_k = \bar{P} > 0$  exists. Then  $\lim_{k \rightarrow \infty} s_k$  exists and is a positive number  $\bar{s} > 0$ .*

**Proof:** As discussed above, the proof is trivial if the trajectory converges in a finite number of sampling intervals or if it converges to the origin. In fact, we only need consider when (9.10) holds for  $t_k$  tending to infinity, or equivalently when  $x_k^T P'_k x_k > x_k^T [Q_k + P_k B_k R_k^{-1} B_k^T P_k] x_k$  for  $k \rightarrow \infty$ . Using (9.10) we may thus write

$$\begin{aligned} s_k &= \frac{x_{k-1}^T P_{k-1} x_{k-1}}{x_{k-1}^T P_k x_{k-1}} s_{k-1} \\ &= \frac{x_{k-1}^T P_{k-1} x_{k-1}}{x_{k-1}^T P_k x_{k-1}} \frac{x_{k-2}^T P_{k-2} x_{k-2}}{x_{k-2}^T P_{k-1} x_{k-2}} \dots \frac{x_0^T P_0 x_0}{x_0^T P_1 x_0} \cdot 1 \\ &= \prod_{i=1}^k \frac{x_{i-1}^T P_{i-1} x_{i-1}}{x_{i-1}^T P_i x_{i-1}} \end{aligned} \quad (9.31)$$

so that as  $k$  tends to infinity, we are trying to guarantee convergence of an infinite product of nonzero (positive) terms. If we now consider the last line of (9.31), for sufficiently small  $\delta$  we may use the analyticity of  $P$  to write  $P_i \approx P_{i-1} + \delta P'_{i-1}$  or equivalently  $P_{i-1} \approx P_i - \delta P'_{i-1}$ . Substituting this into (9.31) and dividing both the numerator and denominator term by term by  $x_{i-1}^T P_i x_{i-1}$ , and letting  $i$  tend to infinity, (9.31) becomes

$$\lim_{k \rightarrow \infty} s_k = \prod_{i=1}^{\infty} (1 + a_i) \quad (9.32)$$

where

$$a_i = -\delta \frac{x_{i-1}^T P'_{i-1} x_{i-1}}{x_{i-1}^T P_i x_{i-1}} \quad (9.33)$$

Now, a well known [1] sufficient condition for convergence of (9.32) to a nonzero value is absolute convergence of the infinite series

$$\sum_{i=1}^{\infty} a_i \quad (9.34)$$

or, equivalently, convergence of

$$\sum_{i=1}^{\infty} |a_i| \quad (9.35)$$

This series is in turn known to converge if it passes the ratio test [1]

$$R = \lim_{i \rightarrow \infty} \sup \left| \frac{a_{i+1}}{a_i} \right| < 1 \quad (9.36)$$

Using (9.33) in (9.36) we thus seek to prove

$$R = \lim_{i \rightarrow \infty} \sup \left| \frac{\frac{x_i^T P'_i x_i}{x_i^T P_{i+1} x_i}}{\frac{x_{i-1}^T P'_{i-1} x_{i-1}}{x_{i-1}^T P_i x_{i-1}}} \right| < 1 \quad (9.37)$$

or

$$\lim_{i \rightarrow \infty} \sup \left| \frac{x_i^T P'_i x_i}{x_i^T P_{i+1} x_i} \right| < \lim_{i \rightarrow \infty} \inf \left| \frac{x_{i-1}^T P'_{i-1} x_{i-1}}{x_{i-1}^T P_i x_{i-1}} \right| \quad (9.38)$$

Now recall that we only need to show (9.38) holds when

$$x_i^T P'_i x_i \geq x_i^T [Q_i + P_i B_i R_i^{-1} B_i^T P_i] x_i \quad (9.39)$$

holds, and we know  $x_i \rightarrow \bar{x} \neq 0$  and  $P_i \rightarrow \bar{P} > 0$ , so that for sufficiently large  $i$ , the right hand side of (9.39) is positive and thus the numerators and denominators of both sides of (9.38) are also positive, allowing the absolute value signs in (9.38) to be removed. Thus  $x_i^T P'_i x_i > 0$  is a positive sequence which we know must converge to zero since the  $P_i$  converge, so that the  $a_i$  satisfy the necessary condition for convergence of an infinite series,  $\lim_{i \rightarrow \infty} a_i = 0$ . Now, since  $P'$  and  $x$  are analytic we see that for sufficiently large  $i$  the values of  $x_i^T P'_i x_i > 0$  fall along a decreasing line going through zero. Thus, we have

$$x_i^T P'_i x_i = x_{i-1}^T P'_{i-1} x_{i-1} - \epsilon_{i-1} \quad (9.40)$$

where  $\epsilon_{i-1}$  is a small positive number. If we again make use of the analyticity of  $P$  and  $x$  to write

$$\begin{aligned} x_i^T P_i x_i &= x_{i-1}^T P_{i-1} x_{i-1} + \delta_i x_{i-1}^T [P'_{i-1} + P_{i-1} F_{i-1} + F_{i-1}^T P_{i-1}] x_{i-1} + x_{i-1}^T \mathcal{O}(\delta_i^2) x_{i-1} \\ &= x_{i-1}^T P_{i-1} x_{i-1} + \delta_i x_{i-1}^T [P'_{i-1} - Q_{i-1} - P_{i-1} B_{i-1} R_{i-1}^{-1} B_{i-1}^T P_{i-1}] x_{i-1} \\ &\quad + x_{i-1}^T \mathcal{O}(\delta_i^2) x_{i-1} \end{aligned} \quad (9.41)$$

then positivity of (9.39) and smallness of  $\delta_i$  allow us to write

$$x_i^T P_i x_i = x_{i-1}^T P_{i-1} x_{i-1} + v_{i-1} \quad (9.42)$$

where  $v_{i-1}$  is again a small positive number. Thus, using (9.40) and (9.42) we see that (9.38) reduces to

$$\frac{x_{i-1}^T P'_{i-1} x_{i-1} - \epsilon_{i-1}}{x_{i-1}^T P_{i-1} x_{i-1} + v_{i-1}} < \frac{x_{i-1}^T P'_{i-1} x_{i-1}}{x_{i-1}^T P_{i-1} x_{i-1}} \quad (9.43)$$

which clearly holds so that the Ratio test is satisfied, and the theorem is proven. ■

We conclude this section by stating that the assumption of  $\lim_{k \rightarrow \infty} s_k \neq 0$  found in all our earlier theorems may thus be replaced by the assumption that  $P_k$  converges to a positive definite matrix limit.

## 9.4 Examples

We now illustrate the preceding theorems by means of some simple examples.

### Example 1

Consider the system

$$\begin{aligned}\dot{x}_1 &= x_1 x_2 + u_1 \\ \dot{x}_2 &= x_1 + u_2 \\ h &= x_1 x_2 + x_2\end{aligned}\tag{9.44}$$

with  $R = I$ . Note that the set of open loop equilibrium points  $A_0$  equals the  $x_2$  axis, i.e, all  $x$  such that  $x_1 = 0$ . Also note that the set where  $y = 0$  equals the union of the  $x_1$  axis and the line  $x_1 = -1$ . Thus, for any  $H$  we choose, there can be no  $x \neq 0$  both in  $A_0$  and the nullspace of  $H$ , so that the conditions of Theorem 9.3.3 are satisfied. If we choose

$$A(x) = \begin{bmatrix} x_2 & 0 \\ 1 & 0 \end{bmatrix}\tag{9.45}$$

and  $H(x) = [x_2 \ 1]$ , then we find

$$\mathcal{O}_f = \begin{bmatrix} H(x) \\ H(x)A(x) \end{bmatrix} = \begin{bmatrix} x_2 & 1 \\ 1 + x_2^2 & 0 \end{bmatrix}\tag{9.46}$$

so that  $\mathcal{O}_f$  is full rank for all  $x$ , and thus  $\mathcal{O}_f^\perp = \{0\}$ . Let us now construct  $\Omega_p$  according to the recursive procedure given in Section 9.3. We have

$$\Omega_0 = \text{rowspan } dh = \text{rowspan } [x_2 \ x_1 + 1]\tag{9.47}$$

$$\Omega_1 = \Omega_0 + dL_a h\tag{9.48}$$

A simple computation gives

$$\Omega_1 = \text{rowspan} \begin{bmatrix} x_2 & x_1 + 1 \\ x_2^2 + 2x_1 + 1 & 2x_1x_2 \end{bmatrix} \quad (9.49)$$

which equals the observable space on at least an open and dense subset of  $\mathcal{R}^2$ . Clearly,  $\Omega_1^\perp$  contains nontrivial elements, for example,  $(x_1, x_2) = (-1, 0)$ . However, since  $\mathcal{O}_f^\perp = \{0\}$  we have satisfied Theorem 9.3.3, and Theorem 9.3.6 tells us  $\mathcal{O}_{nl}^\perp$  is irrelevant and that bounded trajectories can only converge to the origin, so that closed loop stability is guaranteed (provided we sample fast enough and the SDRE solution converges). Additionally we note that for this system  $B = I$  is full rank and constant, which implies the system is controllable for all  $x$  (both in a factored and nonlinear sense), and thus the algorithm is globally well-defined. This example thus shows that the factored observability can play a crucial role in SDRE application. In the next example we illustrate the importance of Theorem 9.3.3.

### Example 2

Consider the system

$$\begin{aligned} \dot{x}_1 &= x_1x_2 + x_2 + u_1 \\ \dot{x}_2 &= x_2^2 + u_2 \\ h &= x_1x_2 + x_2 \end{aligned} \quad (9.50)$$

again with  $R = I$ . The system has the same control input matrix  $B = I$  and output  $h$  as in Example 1, but we have altered the open loop dynamics  $a$ . Now we see that  $Ax = 0$  for all  $x$  such that  $x_2 = 0$ , and  $h = 0$  also at such points. This means that for any choice of  $A$  and  $H$  we cannot satisfy Theorem 9.3.3, and the closed loop system will contain equilibrium points all along the  $x_1$  axis. Thus, even though we have full controllability of this system, the dynamics and chosen output combine in such a way that any trajectory either starting on or passing through the  $x_1$  axis will be fixed at the point of intersection. To illustrate this discussion, choose

$$A = \begin{bmatrix} x_2 & 1 \\ 0 & x_2 \end{bmatrix} \quad (9.51)$$

and  $h$  as in Example 1. Simple computations give

$$\mathcal{O}_f = \begin{bmatrix} x_2 & 1 \\ 0 & x_2 \end{bmatrix} \quad (9.52)$$

and

$$\Omega_1 = \begin{bmatrix} x_2 & x_1 + 1 \\ 2x_2^2 & 4x_2(x_1 + 1) \end{bmatrix} \quad (9.53)$$

where from here on it is implicitly understood that the rowspans of the above distributions are implied by the above notation. As predicted, we see  $\mathcal{O}_f^\perp = \{x \in \mathcal{R}^2 \mid x_2 = 0\}$ , and we also see that  $\Omega_1^\perp = \{x \in \mathcal{R}^2 \mid x_2 = 0 \text{ or } x_1 = -1\}$ . The intersection of the two is thus the  $x_1$  axis as discussed above. In these first two examples we have illustrated the importance of the relationship between  $A$  and  $H$  to observability and hence stability. In the next two examples we also incorporate the effects of the  $B$  matrix, so as to illustrate the effects of controllability/stabilizability as well.

### Example 3

Consider the system

$$\begin{aligned} \dot{x}_1 &= x_1 x_2 + u \\ \dot{x}_2 &= -(1 + x_1^2)x_2 \\ h &= x_1 \end{aligned} \quad (9.54)$$

so that  $H = [1 \ 0]$ , and where  $R = 1$ . Note that the intersection between the open loop equilibrium points (the  $x_1$  axis) and the points in the nullspace of  $H$  (the  $x_2$  axis) consists of only the zero vector, so that Theorem 9.3.3 holds. Also note that  $B = [1 \ 0]^T$  and that for  $x_1 = 0$  we thus clearly do not have full controllability of  $x_2$ . Choosing

$$A = \begin{bmatrix} x_2 & 0 \\ 0 & -(1 + x_1^2) \end{bmatrix} \quad (9.55)$$

we see that  $\{A, B\}$  is stabilizable, but not controllable for all  $x$ . For the given  $H$  we find

$$\mathcal{O}_f = \begin{bmatrix} 1 & 0 \\ x_2 & 0 \end{bmatrix} \quad (9.56)$$

and

$$\Omega_1 = \begin{bmatrix} 1 & 0 \\ x_2 & x_1 \end{bmatrix} \quad (9.57)$$

so that the intersection of  $\mathcal{O}_f^\perp$  and  $\mathcal{O}_{nl}^\perp$  can be no bigger than the  $x_2$  axis. However, along the  $x_2$  axis we see that the dynamics reduce to

$$\dot{x}_2 = -x_2 \quad (9.58)$$

so that by Theorem 9.3.8 we can conclude closed loop stability for sufficiently rapid sampling.

#### Example 4

In this example we illustrate the results of different choices of  $H$  for a given system, incorporating virtually all the theory developed in this chapter. Consider the system

$$\dot{x}_1 = x_1 x_2 + x_2$$

$$\dot{x}_2 = u$$

$$h_1 = [x_1 \ x_2]^T = Hx = Ix \quad (9.59)$$

$$h_2 = x_2 = Hx = [0 \ 1]x \quad (9.60)$$

We can immediately deduce that the set of open loop equilibrium points  $A_0 = \{x \in \mathcal{R}^2 \mid x_1 = -1 \text{ or } x_2 = 0\}$ . Thus, taking  $h = h_1$  we satisfy the requirements of Theorem 9.3.3, but taking  $h = h_2$  we have that the  $x_1$  axis consists entirely of closed loop equilibrium points. Now, it is easily seen that the above system has an invariant, uncontrollable set which is  $\mathcal{C}_{nl}^\perp = \{x \in \mathcal{R}^2 \mid x_1 = -1\}$ , so that any trajectories which start in or enter  $\mathcal{C}_{nl}^\perp$  remain there for all future time. However, if we choose

$$A(x) = \begin{bmatrix} x_2 & 1 \\ 0 & 0 \end{bmatrix} \quad (9.61)$$

then

$$A(x)B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (9.62)$$

which gives the factored controllability matrix

$$M_{cf}(x) = [B \ AB] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (9.63)$$

This matrix is full rank for all  $x$ , so that the system has a controllable factorization for all  $x$ . Now, for  $h = h_1$  we have  $H = I$  so that both  $\mathcal{O}_f = [H^T \ A^T H^T]^T$  and  $\mathcal{O}_{nl} = [H^T \ J^T H^T]^T$ , where  $J$  is the Jacobian matrix of  $a$ , are clearly full rank for all  $x$ . Thus, the only possible closed loop equilibrium point for this choice of regulated output is the origin. However, even though we have a globally controllable factorization, the lack of true controllability for trajectories passing through  $\mathcal{C}_{nl}^\perp$  prevents them from reaching the origin. Thus, by Lasalle's Invariance Principle, we expect any such trajectory to be unbounded. Figure 9.1 shows such a trajectory which starts from  $x_0 = [-1 \ 2]^T$ , and indeed, we see the  $x_2$  state growing unbounded while the  $x_1$  state remains fixed at  $x_1 = -1$ . In Figure 9.2 we give time histories of the Lyapunov function  $V$  and the scaling factor  $s$ . Note that, although  $V$  is continuous and decreasing for all time, the unbounded growth of  $x$  drives  $s$  to zero as time increases.

If, however, we now take  $h = h_2$ , we find

$$\mathcal{O}_f = \mathcal{O}_{nl} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (9.64)$$

so that the  $x_1$  axis is unobservable and receives no control action. Thus, bounded trajectories will converge to the set where  $x_2 = 0$ . Technically, such a choice of  $H$  does not give a detectable factorization, but if we solve the SDRE for only the  $\dot{x}_2$  equation, the control is well-defined. Doing so, we find that trajectories that intersect  $\mathcal{C}_{nl}^\perp$  now converge to the only allowable closed loop equilibrium point,  $x = [-1 \ 0]^T$ . Figures 9.3 and 9.4 are simulation plots illustrating this result. Note that  $s$  is well-behaved in this simulation due to the convergence of  $x_k$  and  $P_k$ , and in fact,  $x^T P x$  would serve equally well as a Lyapunov function in this case since the scaling remains identically one. For both simulations the sample rate was 20 Hz.

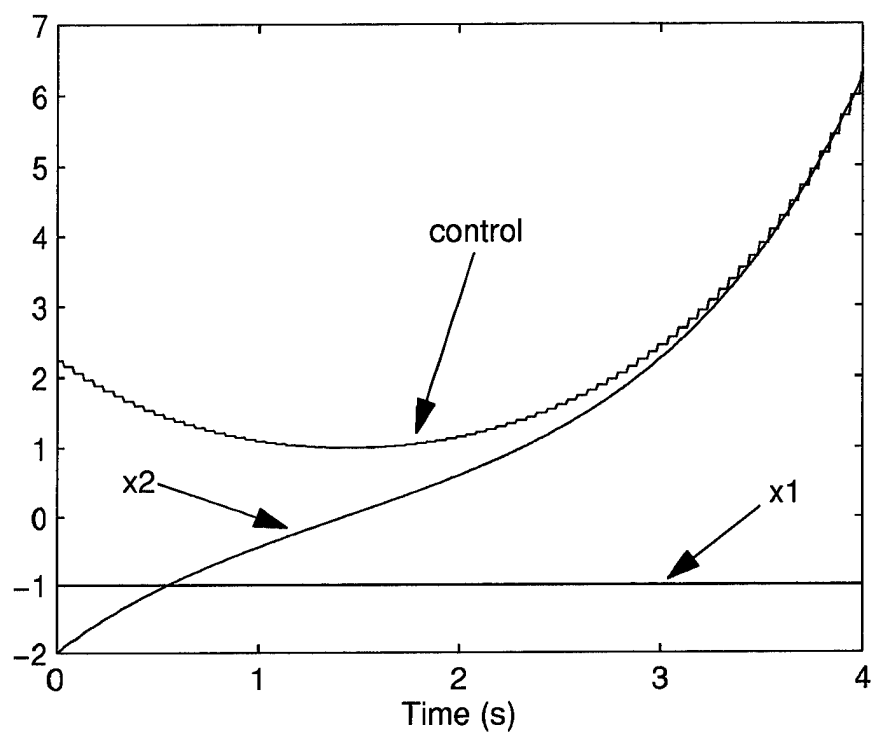


Figure 9.1: State and Control Histories for  $h = h_1$

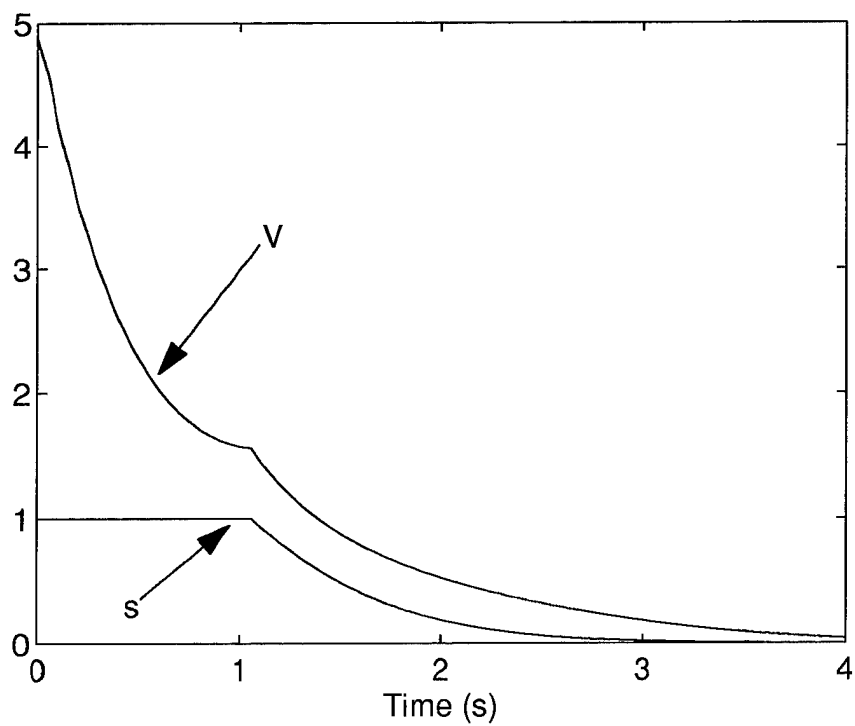


Figure 9.2: Histories of  $V$  and  $s$  for  $h = h_1$

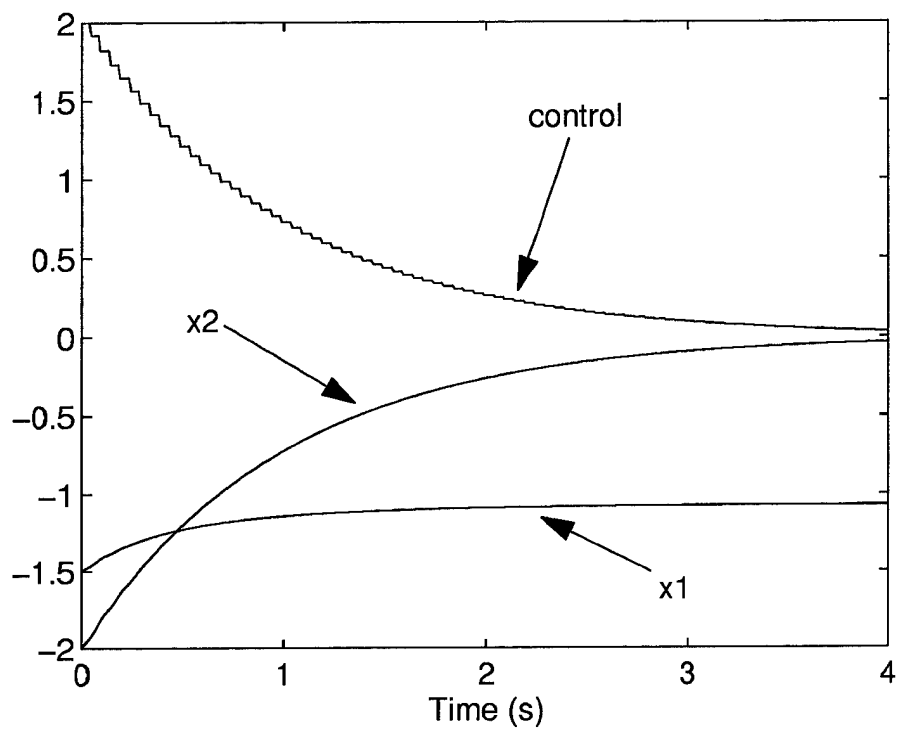


Figure 9.3: State and Control Histories for  $h = h_2$

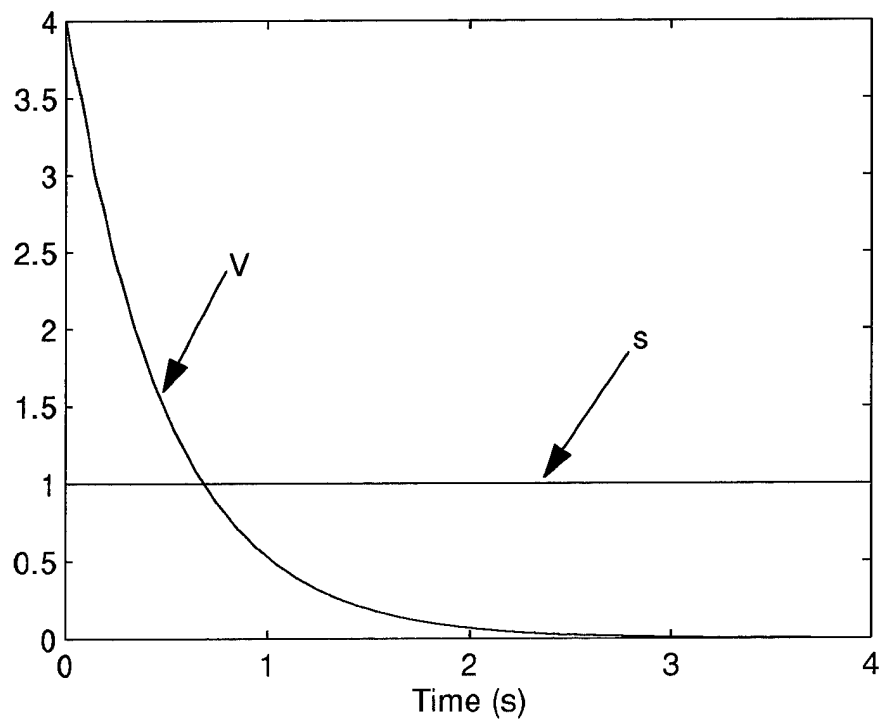


Figure 9.4: Histories of  $V$  and  $s$  for  $h = h_2$

## 9.5 *Summary and Conclusions*

We have shown that it is possible to choose positive semidefinite state weighting matrices in a nonlinear SDRE suboptimal regulation problem and still maintain closed loop stability, provided

- the system is globally analytic and pointwise stabilizable/detectable with respect to  $x$
- the system is sampled and controlled often enough
- the SDRE solution converges

Observability of both the true nonlinear and pointwise linearized systems were seen to play a key role in the determination of invariant sets, while controllability issues were seen to affect convergence of the SDRE solution. We have also established a necessary condition for global asymptotic stability of the closed loop system, which is that the regulated output must not be chosen so that open loop equilibrium points yield a zero output value. We then showed that this condition is automatically satisfied if  $\{H(x), A(x)\}$  is detectable for all  $x$ . The results were derived from a discrete time version of LaSalle's Invariance principle applied to the sample data control law and continuous time plant dynamics. Examples and simulations were given to illustrate and verify the theory.

## *X. Exponential Stability of SDRE Regulators*

### *10.1 Introduction*

In this chapter we investigate potential exponential stability of SDRE regulators. Our approach is to begin by looking at systems whose dynamics can be described by piecewise constant linear representations over fixed intervals, and then extrapolate the results to SDRE regulators by including perturbation terms in the dynamics. Our final result is a conjecture suggesting conditions under which exponential stability of SDRE regulators may be obtained.

### *10.2 Constant $F_k$ Matrices*

We wish to consider stability of the nonlinear autonomous dynamical system

$$\dot{x}(t) = f(x) = F(x(t))x(t) \quad (10.1)$$

with state vector  $x \in \mathcal{R}^n$ , and with  $F(x(t))$  defined to be equal to a sequence of Hurwitz matrices,  $F_k$ , which are assumed constant over finite time intervals  $[t_k, t_{k+1})$ . That is,

$$\dot{x}(t) = F_k x(t) \quad \forall \quad t \in [t_k, t_{k+1}) \quad (10.2)$$

and  $\text{Re}[\lambda_i(F_k)] < 0 \quad \forall \quad i, k$ . Note that for all practical purposes we could thus consider  $F(x(t))$  to be a function of time only, so that (10.1) would be linear time-varying, but we maintain the  $x$  dependency for later use. For such a system we may write with no error

$$x_{k+1} = \Phi_k x_k \quad (10.3)$$

where  $\Phi_k = e^{F_k(t_{k+1}-t_k)}$  is the state transition matrix from  $t_k$  to  $t_{k+1}$ . Let us define  $\delta_k = t_{k+1} - t_k$  and now choose any initial condition  $x(t_0) = x_0 \in \mathcal{R}^n$ . We seek to discover conditions under which the norm of the state vector will eventually converge to zero, i.e., the origin is attractive. To this end we consider the value of  $x_1$ , which from (10.3) may be written as

$$x_1 = \Phi_0 x_0 \quad (10.4)$$

Defining  $\|x\|$  to be the vector 2-norm (Euclidean norm), we may write

$$\|x_1\| \leq \bar{\sigma}(\Phi_0)\|x_0\| \quad (10.5)$$

where  $\bar{\sigma}$  represents the maximum singular value. Now, we may invoke the definition of  $\Phi$  and the properties of the exponential of a Hurwitz matrix to write [42]

$$\bar{\sigma}(\Phi_0) \leq c_0 e^{-d_0 \delta_0} \quad (10.6)$$

for constants  $d_0 > 0$  and  $c_0 \geq 1$ , so that (10.5) becomes

$$\|x_1\| \leq c_0 e^{-d_0 \delta_0} \|x_0\| \quad (10.7)$$

and we see that the norm of  $x$  is not guaranteed to be decreasing over the first time interval since  $c_0 \geq 1$ . If we now consider the value of the state vector at  $t_2$  we have

$$x_2 = \Phi_1 \Phi_0 x_0 \quad (10.8)$$

so that by defining  $d_1$ ,  $c_1$  and  $\delta_1$  for  $\Phi_1$  as in (10.6) we have

$$\|x_2\| \leq c_0 c_1 e^{-d_1 \delta_1} e^{-d_0 \delta_0} \|x_0\| \quad (10.9)$$

If we now let  $\Delta_2 = t_2 - t_0 = \delta_1 + \delta_0$  it is clear that we can write

$$\|x_2\| \leq c_0 c_1 e^{-m_2 \Delta_2} \|x_0\| \quad (10.10)$$

for some  $m_2 > 0$  ( $m_2 = \frac{1}{2}(d_1 + d_0)$  if  $\delta_1 = \delta_0$ ). By induction, after  $r$  time steps we can therefore write

$$\|x_r\| \leq c_0 c_1 \dots c_r e^{-m_r \Delta_r} \|x_0\| \quad (10.11)$$

Now if we take the limit of (10.11) as  $r \rightarrow \infty$ , the exponential term tends to 0 provided  $m_r$  does not approach 0, but we are not able to conclude that  $\|x_r\| \rightarrow 0$  because the infinite product of the  $c_k$ 's may grow faster than the exponential term decays. This is why simply guaranteeing negative eigenvalues in a closed loop matrix for a linear time-varying system is insufficient to conclude stability. Negative eigenvalues of the  $F_k$  guarantee eigenvalues of  $\Phi_k$  inside the unit disk, but not

that  $\|\Phi_k\| < 1$ . Thus, we shall return to (10.8) and seek additional conditions on the  $F_k$  beside being Hurwitz which lead to attractiveness of the origin. By using the Jordan form of  $F_k = M_k J_k M_k^{-1}$ , we may write (10.8) as

$$x_2 = e^{F_1 \delta_1} e^{F_0 \delta_0} x_0 = M_1 e^{J_1 \delta_1} M_1^{-1} M_0 e^{J_0 \delta_0} M_0^{-1} x_0 \quad (10.12)$$

Now, if we define  $\Delta M_1 \equiv M_0 - M_1$ , then we can write

$$\begin{aligned} x_2 &= M_1 e^{J_1 \delta_1} M_1^{-1} (M_1 + \Delta M_1) e^{J_0 \delta_0} M_0^{-1} x_0 \\ &= M_1 e^{J_1 \delta_1} e^{J_0 \delta_0} M_0^{-1} x_0 + M_1 e^{J_1 \delta_1} M_1^{-1} \Delta M_1 e^{J_0 \delta_0} M_0^{-1} x_0 \end{aligned} \quad (10.13)$$

At this point we assume each  $F_k$  is diagonalizable (simple) so that  $J_k = D_k$ , a diagonal matrix of the eigenvalues of  $F_k$ , all of which have negative real part by the Hurwitz assumption. Equation (10.13) thus becomes

$$\begin{aligned} x_2 &= M_1 e^{D_1 \delta_1} e^{D_0 \delta_0} M_0^{-1} x_0 + M_1 e^{D_1 \delta_1} M_1^{-1} \Delta M_1 e^{D_0 \delta_0} M_0^{-1} x_0 \\ &= M_1 e^{D_1 \delta_1 + D_0 \delta_0} M_0^{-1} x_0 + M_1 e^{D_1 \delta_1} M_1^{-1} \Delta M_1 e^{D_0 \delta_0} M_0^{-1} x_0 \\ &= M_1 e^{\bar{D}_2 \Delta_2} M_0^{-1} x_0 + M_1 e^{D_1 \delta_1} M_1^{-1} \Delta M_1 e^{D_0 \delta_0} M_0^{-1} x_0 \end{aligned} \quad (10.14)$$

where  $\bar{D}_2$  is a diagonal matrix of possibly complex numbers, all of which have negative real part. In the event  $\delta_1 = \delta_0$ , we have  $\bar{D}_2 = \frac{1}{2}(D_1 + D_0)$ , so that  $\bar{D}_2$  is just the average of the  $D_k$ . If we now take the norm of (10.14), use the triangle inequality and the submultiplicative property, and invoke some singular value properties we find

$$\|x_2\| \leq \frac{\bar{\sigma}(M_1)}{\underline{\sigma}(M_0)} e^{-m_{20} \Delta_2} \|x_0\| + \frac{\bar{\sigma}(M_1)}{\underline{\sigma}(M_0)} e^{d_1 \delta_1} \frac{\bar{\sigma}(\Delta M_1)}{\underline{\sigma}(M_1)} e^{d_0 \delta_0} \|x_0\| \quad (10.15)$$

where

$$m_{20} \equiv \max_i |Re[\lambda_i(\bar{D}_2)]| > 0 \quad (10.16)$$

and  $d_1$  and  $d_0$  are defined likewise for  $D_1$  and  $D_0$ , respectively. Suppose now that we define a scalar valued function that overbounds the error term due to the change in  $M_k$ , so that

$$\frac{\bar{\sigma}(\Delta M_k)}{\underline{\sigma}(M_k)} \leq \eta(k) \quad (10.17)$$

where  $\eta(k)$  is simply a function of  $k$  unspecified as of yet, and we define  $\eta(0) = 1$ . Then (10.15) becomes

$$\|x_2\| \leq \frac{\bar{\sigma}(M_1)}{\underline{\sigma}(M_0)} e^{-m_{20}\Delta_2} \|x_0\| + \frac{\bar{\sigma}(M_1)}{\underline{\sigma}(M_0)} \eta(1) e^{-m_{21}\Delta_2} \|x_0\| \quad (10.18)$$

for some  $m_{21} > 0$ . Recall now that the columns of  $M_k$  are just the generalized eigenvectors of  $F_k$ , and thus may be chosen to have any desired norm. If we choose the  $M_k$  such that each column of  $M_k$  has unit norm, then it is trivial to show that  $\bar{\sigma}(M_k) \leq \sqrt{n}$  for any  $k$ , where  $n$  is the dimension of the state vector  $x$ . Using this fact we may define the constant

$$K = \frac{\sqrt{n}}{\underline{\sigma}(M_0)} \quad (10.19)$$

so that (10.18) may be written

$$\|x_2\| \leq K e^{-m_{20}\Delta_2} \|x_0\| + K \eta(1) e^{-m_{21}\Delta_2} \|x_0\| \quad (10.20)$$

Proceeding in a similar manner it is easily established that

$$\begin{aligned} \|x_3\| &\leq K e^{-m_{30}\Delta_2} \|x_0\| + K \eta(1) e^{-m_{311}\Delta_2} \|x_0\| \\ &\quad + K \eta(2) e^{-m_{312}\Delta_2} \|x_0\| + K \eta(1) \eta(2) e^{-m_{32}\Delta_2} \|x_0\| \end{aligned} \quad (10.21)$$

for some positive constants  $m_{30}$ ,  $m_{311}$ ,  $m_{312}$  and  $m_{32}$ . Grouping terms of like powers of  $\eta$  we get

$$\|x_3\| \leq K e^{-m_{30}\Delta_2} \|x_0\| + 2K \mathcal{O}(\eta) e^{-m_{31}\Delta_2} \|x_0\| + K \mathcal{O}(\eta^2) e^{-m_{32}\Delta_2} \|x_0\| \quad (10.22)$$

for some positive constant  $m_{31}$ . By induction it can thus be shown that

$$\begin{aligned} \|x_r\| &\leq K a_{r0} e^{-m_{r0}\Delta_r} \|x_0\| + K a_{r1} \mathcal{O}(\eta) e^{-m_{r1}\Delta_r} \|x_0\| \\ &\quad + K a_{r2} \mathcal{O}(\eta^2) e^{-m_{r2}\Delta_r} \|x_0\| + \dots \\ &\quad + K a_{r(r-1)} \mathcal{O}(\eta^{(r-1)}) e^{-m_{r(r-1)}\Delta_r} \|x_0\| \end{aligned} \quad (10.23)$$

for positive constants  $m_{rj}$  and  $a_{rj}$ , where  $a_{rj}$  is given by the shifted binomial coefficient formula

$$a_{rj} = \binom{r-1}{j} = \frac{(r-1)!}{j!(r-j-1)!} \quad (10.24)$$

If we assume all of the  $\eta(k)$  are bounded above by a positive constant  $\rho$ , then from (10.23) we may write

$$\|x_r\| \leq K\|x_0\| \sum_{j=0}^{r-1} a_{rj} \rho^j e^{-m_{rj} \Delta_r} \quad (10.25)$$

If we now take

$$m_r = \min_j m_{rj} > 0 \quad (10.26)$$

we may pull the exponential term out of the sum in (10.25) to get

$$\|x_r\| \leq K\|x_0\| e^{-m_r \Delta_r} \sum_{j=0}^{r-1} a_{rj} \rho^j \quad (10.27)$$

Now, recognizing that

$$\sum_{j=0}^{r-1} a_{rj} \rho^j = (1 + \rho)^r \quad (10.28)$$

is just a special case of the binomial formula, we have

$$\|x_r\| \leq K\|x_0\| e^{-m_r \Delta_r} (1 + \rho)^r \quad (10.29)$$

Now, define

$$\bar{\delta}_r \equiv \frac{1}{r} \sum_{j=0}^{r-1} \delta_j = \frac{1}{r} \Delta_r \quad (10.30)$$

so that  $\bar{\delta}_r$  is just the average sampling interval size. Then (10.29) becomes

$$\|x_r\| \leq K\|x_0\| \left( \frac{1 + \rho}{e^{m_r \bar{\delta}_r}} \right)^r \quad (10.31)$$

and thus it can easily be seen that for the norm of  $x$  to decrease as the number of time steps  $r$  increases, we must have

$$\frac{1 + \rho}{e^{m_r \bar{\delta}_r}} < 1 \quad (10.32)$$

Taking the natural log of both sides of (10.32) and rearranging, we get the equivalent condition

$$m_r \bar{\delta}_r > \ln(1 + \rho) \quad (10.33)$$

We now make one further assumption on the original system (10.1). If  $f$  is continuously differentiable, then in some finite neighborhood of the origin the system dynamics are governed by the linearization

of  $f$ , so that  $F(x) = \frac{\partial f}{\partial x}|_{x=0}$  for small  $\|x\|$ . Thus, near the origin not only is  $F$  Hurwitz by our original assumption, but it is also constant. The linearized dynamics  $\dot{x} = F(0)x$  have some finite, bounded domain of attraction,  $\mathcal{D}_l$ , which is an invariant set and for which every  $x \in \mathcal{D}_l$  tends to 0 as  $r \rightarrow \infty$ . Define  $\mathcal{B}_1$  as the radius of the largest ball in  $\mathcal{R}^n$  (defined by the vector 2-norm) that is wholly contained in  $\mathcal{D}_l$ , and define  $\mathcal{B}_2$  as the radius of the smallest ball wholly containing  $\mathcal{D}_l$ . Then, if (10.33) is satisfied, it follows that any trajectory of (10.1) is bounded and converges to the origin. This may be seen by considering two cases corresponding to trajectories which start either inside or outside of  $\mathcal{D}_l$ . If  $\|x_0\| \leq \mathcal{B}_1$  so that the initial condition lies inside the domain of attraction of the linearized dynamics, then clearly  $\|x_r\| \leq \mathcal{B}_2 \forall r$  since  $\mathcal{D}_l$  is an invariant set, and additionally we have  $x_r \rightarrow 0$  as  $r \rightarrow \infty$  by definition of  $\mathcal{D}_l$ . If on the other hand  $\|x_0\| \geq \mathcal{B}_1$ , the desired properties follow because if (10.33) is satisfied, every system trajectory that starts outside of  $\mathcal{D}_l$  enters  $\mathcal{D}_l$  after a finite number of time steps  $l$ , which is computable by setting

$$\|x_l\| \leq K \|x_0\| \left( \frac{1+\rho}{e^{m_l \bar{\delta}_l}} \right)^l \leq \mathcal{B}_1 \quad (10.34)$$

and solving for  $l$  to get

$$l \geq \frac{\ln \left( \frac{\mathcal{B}_1}{K \|x_0\|} \right)}{\ln(1+\rho) - m_l \bar{\delta}_l} \quad (10.35)$$

Boundedness follows since outside of  $\mathcal{D}_l$ , we have from (10.31) and (10.32) that for any  $r$

$$\|x_r\| \leq K \|x_0\| \quad (10.36)$$

while inside  $\mathcal{D}_l$  we have  $\|x_r\| \leq \mathcal{B}_2$ , giving the overall bound

$$\|x_r\| \leq \max\{K \|x_0\|, \mathcal{B}_2\} \quad (10.37)$$

Convergence of  $x$  to 0 follows trivially from  $x_l$  being in  $\mathcal{D}_l$  after  $l$  time steps. Note that in the above development we allowed variable sampling intervals, and we ended up using the minimum of all the  $m_{r,j}$  as our exponential decay factor. If we instead fix  $\delta_j = \bar{\delta}_r$  constant, then the above development also holds for

$$m_r = \frac{1}{r} \sum_{j=0}^{r-1} m_{r,j} \quad (10.38)$$

so that now we are using the average of the  $m_{rj}$ , which could be significantly larger than their minimum, thus making (10.32) or equivalently (10.33) easier to satisfy.

Now, it remains to be seen under what conditions (10.33) can be satisfied. Intuitively, (10.33) is easier to satisfy if  $m_r$  and  $\bar{\delta}_r$  are large, and if  $\rho$  is small. Indeed, if  $\rho = 0$  we simply need  $m_r$  and  $\bar{\delta}_r$  greater than zero, so that we recover the stability condition of a linear time-invariant system as a special case. For nonzero  $\rho$ , however, we have a minimum size requirement that we must meet on the product  $m_r \bar{\delta}_r$ . For systems meeting the analyticity and diagonalizability assumptions on  $F$ , it is known [38] that each column of  $M$  is a linearly independent eigenvector of  $F$  which is analytic with respect to  $t$ . Thus, the size of  $\eta_k$  and hence  $\rho$  is directly proportional to the size of  $\delta_k$ , for sufficiently small  $\delta_k$ . Additionally, we know from Chapter 7 that if we choose  $Q > 0$  globally, then we can increase the size of  $m_r$  for the controllable modes by increasing the minimum eigenvalue of  $Q$ . Thus,  $\delta_k$  and  $Q$  can be used directly to ensure satisfaction of (10.33), provided the uncontrollable modes also have sufficiently small eigenvalues. If  $F$  is not globally simple, then it may be possible to perturb  $Q$  so as to give a globally simple  $F$ . If  $F$  cannot be made simple, the above analysis becomes significantly more difficult, due to lack of guaranteed analyticity of all the eigenvectors of  $F$  [42]. It still may be possible to extend the above analysis, however, by considering the fact that, for sufficiently fast sampling and sufficiently large  $m_r$ , divergent trajectories can only occur along infinite paths where  $F$  is not simple. This is because a trajectory encountering a finite number of such points can be thought of as having a larger premultiplying constant  $K$ , which does not affect whether the system converges or not, but instead only affects time to converge. This line of analysis may be made rigorous by applying the sufficient conditions for convergence of an infinite product [1]. If no such infinite paths exist, or if it can be shown that closed loop trajectories cannot follow such paths, then exponential stability may still be provable. Experimental analysis along these lines supports this conjecture. We now seek to extend the above concepts to exponential stability of the sampled data SDRE nonlinear regulator algorithm.

### 10.3 Exponential Stability of SDRE

We now repose the above stability analysis in the context of nonlinear SDRE regulation, showing that the SDRE problem reduces to a perturbed version of the above problem. We are thus able to show that basically the same conclusions given above apply, given our standard assumptions. We wish to consider stability of the closed loop system resulting from the nonlinear SDRE regulator, given by

$$\dot{x}(t) = f(x(t)) = F(x(t))x(t) \quad (10.39)$$

where  $F(x(t_k)) = F_k$  and for all  $t \in [t_k, t_{k+1})$  we have  $F(x(t)) = F_k + \Delta F_k(x(t))$ . Under our standard assumptions of global analyticity, stabilizability, and detectability of the open loop system matrices, we have  $F$  analytic with respect to  $t$  and  $F_k$  Hurwitz for all  $k$ . Using the above expression for  $F(x(t))$  we have for any  $t \in [t_k, t_{k+1})$

$$\dot{x}(t) = F_k x(t) + \Delta F_k(x(t))x(t) \quad (10.40)$$

which can be seen to be a perturbed version of (10.2), with the error term  $\Delta F_k(x(t))$  representing the effects of the changing nonlinear dynamics over a sampling interval. Thus, starting from any initial condition  $x_0 \in \mathcal{R}^n$  we may write

$$x_1 = e^{F_0 \delta_0} x_0 + \int_0^{t_1} e^{F_0(t_1-\tau)} \Delta F_0(\tau) x(\tau) d\tau \quad (10.41)$$

and we now must be concerned with the effects of the integral error term on the right hand side of (10.41), since we have already derived conditions under which the norm of the terms deriving from the first term will tend to zero. Taking norms we may write

$$\|x_1\| \leq c_0 e^{-m_0 \delta_0} + \left\| \int_0^{t_1} e^{F_0(t_1-\tau)} \Delta F_0(\tau) x(\tau) d\tau \right\| \quad (10.42)$$

If we assume that we are sampling the system fast enough to invoke dominance of the linear terms of the power series expansions of  $\Delta F$  and  $x$ , we may write

$$\Delta F_0(\tau) = F(\tau) - F_0 = F_0 + \tau F'_0 - F_0 = \tau F'_0 \quad (10.43)$$

where  $F'_0$  is the matrix of time derivatives of  $F(x(t))$  evaluated at  $t_0 = 0$ , and

$$x(\tau) = x_0 + \tau F_0 x_0 \quad (10.44)$$

so that the second term on the right hand side of (10.42) gives

$$\begin{aligned} \left\| \int_0^{t_1} e^{F_0(t_1-\tau)} \Delta F_0(\tau) x(\tau) d\tau \right\| &\leq \left\| \int_0^{t_1} \tau e^{F_0(t_1-\tau)} F'_0 x_0 d\tau \right\| \\ &+ \left\| \int_0^{t_1} \tau^2 e^{F_0(t_1-\tau)} F'_0 F_0 x_0 d\tau \right\| \end{aligned} \quad (10.45)$$

$$= \|v0\| + \|w0\| \quad (10.46)$$

where  $v0$  and  $w0$  are vectors respectively corresponding to the values of the integrals of the first and second terms on the right hand side of (10.45), obtained by applying the Mean Value Theorem for integrals [1]. The  $i$ th elements of  $v0$  and  $w0$  are thus given by

$$v0_i = T_i \delta_0 \tau_{01i} e^{F_0(t_1-\tau_{01i})} F'_0 x_0 \quad (10.47)$$

and

$$w0_i = T_i \delta_0 \tau_{02i}^2 e^{F_0(t_1-\tau_{02i})} F'_0 F_0 x_0 \quad (10.48)$$

where  $T_i$  is an  $n \times n$  matrix with  $i$ th diagonal element equal to 1 and all other elements equal to 0 ( $T_i$  simply picks off the  $i$ th element of the vector it premultiplies). The  $\tau_{0si} \in [0, \delta_0]$  are the times at which evaluation of the integrands multiplied by the time interval  $\delta_0$  yields the value of the  $i$ th elements of the integrals, for  $s = 1, 2$  respectively corresponding to the first and second integrals on the right hand side of (10.45). If we again assume  $F_k$  is simple, we get

$$\|v0\| \leq K \sqrt{n} \delta_0 \tau_{01} e^{-h_0(t_1-\tau_{01})} \bar{\sigma}(F'_0) \|x_0\| \quad (10.49)$$

and

$$\|w0\| \leq K \sqrt{n} \delta_0 \tau_{02}^2 e^{-h_0(t_1-\tau_{02})} \bar{\sigma}(F'_0) \bar{\sigma}(F_0) \|x_0\| \quad (10.50)$$

where  $h_0 > 0$  is the absolute value of the maximum real part of the eigenvalues of  $F_0$ ,  $K = \frac{\bar{\sigma}(M_0)}{\underline{\sigma}(M_0)} > 0$  as before, and  $\tau_{01}$  and  $\tau_{02}$  satisfy

$$\tau_{01} = \arg \max_i \tau_{01i} e^{-h_0(t_1-\tau_{01i})} \quad (10.51)$$

$$\tau_{02} = \arg \max_i \tau_{02i}^2 e^{-h_0(t_1 - \tau_{02i})} \quad (10.52)$$

Note that if  $\tau_{0s} = 0$  then  $\tau_{0si} = 0$  for all  $i$  and the norm of  $\|v0\|$  (for  $s = 1$ ) or  $\|w0\|$  (for  $s = 2$ ) is zero. We thus assume  $\tau_{0s} > 0$  and since  $\tau_{0s} \leq \delta_0$  we may write

$$\|v0\| \leq K\sqrt{n}\delta_0^2 e^{-h_0(t_1 - \tau_{01})} \overline{\sigma}(F'_0) \|x_0\| \quad (10.53)$$

and

$$\|w0\| \leq K\sqrt{n}\delta_0^3 e^{-h_0(t_1 - \tau_{02})} \overline{\sigma}(F'_0) \overline{\sigma}(F_0) \|x_0\| \quad (10.54)$$

As in our previous analysis, we expect the exponential decay present in (10.53) and (10.54) to be crucial to proving stability. We therefore want to ensure that  $\tau_{0s} \neq t_1$ . It is easily verified (by Taylor series expansions for example) that for  $\delta_0$  small enough, the functions  $T_i \tau e^{-F\tau} x$  and  $T_i \tau^2 e^{-F\tau} x$  are increasing functions of  $\tau$  for a fixed Hurwitz matrix  $F$  and fixed vector  $x$ . Since these functions are strictly increasing, by purely geometric arguments it is clear their integrals over small  $\delta_0$  cannot be equal to the product of  $\delta_0$  and the integrand evaluated at the right endpoint of the interval. Thus, by sampling 'fast enough', we can guarantee that  $t_1 - \tau_{0s} > 0$  so that we may in fact write

$$t_1 - \tau_{0s} = \frac{\delta_0}{u_{0s}}, \quad s = 1, 2 \quad (10.55)$$

for some  $u_{0s} \geq 1$ , which combined with (10.53), (10.54), and (10.46) gives

$$\begin{aligned} \left\| \int_0^{t_1} e^{F_0(t_1 - \tau)} \Delta F_0(\tau) x(\tau) d\tau \right\| &\leq K\sqrt{n}\delta_0^2 e^{-a_{01}\delta_0} \overline{\sigma}(F'_0) \|x_0\| \\ &\quad + K\sqrt{n}\delta_0^3 e^{-a_{02}\delta_0} \overline{\sigma}(F'_0) \overline{\sigma}(F_0) \|x_0\| \end{aligned} \quad (10.56)$$

where  $h_0 \geq a_{0s} = \frac{h_0}{u_{0s}} > 0$ ,  $s = 1, 2$ . Since  $a_{0s} \leq h_0$ , we see that these perturbation terms decrease the corresponding effective exponential decay rate. The remainder of the analysis proceeds as when  $\Delta F$  was equal to zero in the first part of this chapter. We assume  $\delta_k$  is selected small enough to guarantee the constant overbounds

$$\delta_k \overline{\sigma}(F'_k) \leq \alpha \quad (10.57)$$

$$\delta_k^2 \overline{\sigma}(F'_k) \overline{\sigma}(F_k) \leq \beta \quad (10.58)$$

so that  $v_0$  contributes a perturbation of order  $\delta$ , and  $w_0$  contributes a perturbation of order  $\delta^2$ . Note that these bounds are only achievable if  $F$  and its derivative are bounded above. Upon expanding terms and retaining only zeroth and first-order terms in the assumed constant  $\bar{\delta}_r = \delta$ , we obtain an expression of the form

$$\|x_r\| \leq K\|x_0\| \left( \frac{1+\rho}{e^{m_r\delta}} \right)^r + K_1\|x_0\| \left( \frac{1+\delta\rho}{e^{a_r\delta}} \right)^r \quad (10.59)$$

where  $K_1 = \alpha\sqrt{n}K > 0$  which, as expected, is the same result as in Section 10.2 plus a first-order perturbation term in  $\delta$ . We thus now gain the attractiveness condition

$$\frac{1+\delta\rho}{e^{a_r\delta}} < 1 \quad (10.60)$$

in addition to (10.32). While the  $\delta$  premultiplying  $\rho$  in the numerator of (10.60) makes (10.60) easier to satisfy than (10.32), we see that we now also have the reduced exponential decay factor  $a_r$  in the denominator, which makes (10.60) harder to satisfy. As discussed above, this decay factor will be smaller than the unperturbed decay factor  $m_r$  in (10.32), but will be no smaller than one half  $m_r$  if we sample rapidly enough to have  $\tau e^{F\tau}$  be effectively linear in  $\tau$  over the sampling interval. If this is the case and (10.32) is satisfied, then (10.60) will also be satisfied provided

$$\rho < (1 - 2\delta)/\delta^2 \quad (10.61)$$

which should be easily satisfied for reasonably small  $\delta$ . Although we have performed only a first-order perturbation analysis, we nevertheless see that exponential stability of SDRE regulators is potentially achievable under the same conditions as given in Section 10.2, plus some additional boundedness conditions on  $F$  and its derivative.

From the analysis performed in this chapter, we see that exponential stability of SDRE regulators may indeed be provable. The line of inquiry pursued here indicates that sufficiently rapid sampling, analyticity of the system matrices, and diagonalizability of the closed loop dynamics matrix play key roles, as well as the standard assumptions of globally stabilizable and detectable factorizations. Difficulty in guaranteeing the diagonalizability condition poses the most serious challenge to proving exponential stability in this way.

## *XI. Extensions to SDRE Nonlinear $H_\infty$ Control*

In this chapter we extend the theory developed in the preceding chapters for nonlinear regulation via the SDRE method to SDRE nonlinear  $H_\infty$  control. We emphasize in particular the similarities and differences between the required assumptions and corresponding theoretical developments. Prior to these developments, however, we give some basic theorems addressing properties of standard and  $H_\infty$  Riccati equations.

### *11.1 Standard Riccati Theory*

Throughout this dissertation, we have relied on stabilizable and detectable system factorizations guaranteeing the existence of positive semidefinite stabilizing solutions to algebraic Riccati equations of the form we encounter when applying SDRE nonlinear regulation. When considering  $H_\infty$  type Riccati equations, we lose this guaranteed solution existence property, except under special circumstances. We thus now give the theorems which explicitly state conditions for existence of ARE solutions, allowing us to see how things change when we consider the SDRE nonlinear  $H_\infty$  control problem. We precede the theorems with some necessary terminology.

Consider the ARE

$$A^T P + PA + PKP + Q = 0 \quad (11.1)$$

where  $A$ ,  $Q$ , and  $K$  are real  $n \times n$  matrices and  $K$  and  $Q$  are symmetric, and the associated Hamiltonian matrix

$$\mathcal{H} \equiv \begin{bmatrix} A & K \\ -Q & -A^T \end{bmatrix} \quad (11.2)$$

We are interested in existence of stabilizing solutions of (11.1), where by stabilizing we mean  $A + KP$  is a Hurwitz matrix. Under certain conditions on  $\mathcal{H}$  (namely the well-known stability and complementarity conditions) [76], a stabilizing solution  $P$  to (11.1) exists, which is uniquely determined by  $\mathcal{H}$ , so that the mapping  $\mathcal{H} \rightarrow P$  is a function. We will denote this function by  $Ric$ , so that

$P = Ric(\mathcal{H})$  if  $\mathcal{H}$  possesses the above two properties, or equivalently, if  $\mathcal{H}$  belongs to the domain of  $Ric$ , which we denote by  $dom(Ric)$ . Thus we may replace the rather verbose expression, ‘If the Hamiltonian matrix  $\mathcal{H}$  possesses the stability and complementarity properties, then there exists a stabilizing solution  $P$  to (11.1)’, with ‘If  $\mathcal{H} \in dom(Ric)$ , then  $P = Ric(\mathcal{H})$ ’. We now give the two theorems on which the SDRE theory of the previous chapters of this dissertation has relied, taken from [76].

**Theorem 11.1.1** *Suppose  $\mathcal{H} \in dom(Ric)$  and  $P = Ric(\mathcal{H})$ . Then*

- i.  $P$  is real symmetric*
- ii.  $P$  satisfies the ARE (11.1)*
- iii.  $A + KP$  is stable*

**Proof:** See [76], Theorem 13.5. ■

**Theorem 11.1.2** *Suppose  $\mathcal{H}$  has the form*

$$\mathcal{H} \equiv \begin{bmatrix} A & -BB^T \\ -H^T H & -A^T \end{bmatrix} \quad (11.3)$$

*Then  $\mathcal{H} \in dom(Ric)$  iff  $\{A, B\}$  is stabilizable and  $\{H, A\}$  has no unobservable modes on the imaginary axis. Furthermore,  $P = Ric(\mathcal{H}) \geq 0$ , and  $Ker(P) = 0$  iff  $\{H, A\}$  has no stable unobservable modes.*

**Proof:** See [76], Theorem 13.7. ■

Note that from Theorem 11.1.2, detectability of  $\{H, A\}$  is sufficient but not necessary for existence of positive semidefinite, stabilizing solutions to (11.1). Also note from the theorem the assumption that the (1,2)-block of  $\mathcal{H}$  be negative semidefinite, which implies the stabilizing solution is maximal, i.e., the stabilizing solution  $P_+$  is such that

$$P_+ - P \geq 0, \forall P \quad (11.4)$$

satisfying (11.1). We note that this negative semidefiniteness assumption on  $K$  does not necessarily hold when we consider the  $H_\infty$  Riccati equation

$$A^T P + PA + P\left(\frac{1}{\gamma^2} GG^T - BB^T\right)P + H^T H = 0 \quad (11.5)$$

and its associated Hamiltonian matrix

$$\mathcal{H} \equiv \begin{bmatrix} A & \frac{1}{\gamma^2} GG^T - BB^T \\ -H^T H & -A^T \end{bmatrix} \quad (11.6)$$

Indeed, it is precisely this lack of sign definiteness which prevents guaranteeing existence of stabilizing solutions to (11.5), so that additional assumptions are required. Upon occasion, we may have that  $K = (1/\gamma^2)GG^T - BB^T \geq 0$ . In this case the following theorem may allow us to conclude existence of stabilizing solutions.

**Theorem 11.1.3** *Define the quadratic matrix function of  $P$*

$$\mathcal{Q}(P) \equiv A^T P + PA + PKP + Q \quad (11.7)$$

*and assume that  $K \geq 0$  and that  $\exists$  a symmetric matrix  $P = P^T$  such that  $\mathcal{Q}(P) \leq 0$ . If  $\{A, K\}$  is stabilizable, then  $\exists$  a unique minimal solution  $P_-$  to (11.1). Furthermore,*

$$P_- \leq P, \forall P \text{ such that } \mathcal{Q}(P) \leq 0 \quad (11.8)$$

*and  $A + KP_-$  has all its eigenvalues in the closed left-half plane. If  $\mathcal{Q}(P) < 0$ , then  $P_- < P$  in (11.8), and  $A + KP_-$  is Hurwitz.*

**Proof:** See [76], Theorem 13.11. ■

Thus, when  $K$  is positive semidefinite, we seek minimal, as opposed to maximal solutions to the ARE.

From these theorems we see that, in general, we shall have to assume existence of stabilizing solutions to (11.5), although for  $K$  either positive or negative semidefinite, and  $\{A, K\}$  stabilizable, we may be guaranteed existence of solutions. With this concept of conditional existence of solutions for (11.5) in hand, we now revisit some of the analysis performed for the regulator type Riccati equations, in the context of  $H_\infty$  type Riccati equations.

## 11.2 Solution Properties of $H_\infty$ Scalar Analytic Systems

In this section we consider SDRE nonlinear  $H_\infty$  control for the scalar (single-state) analytic case, giving necessary and sufficient conditions for obtaining (locally) stabilizing solutions, just as we did for the SDRE nonlinear regulator in Chapter 5. As we did then, we consider control of input-affine nonlinear dynamical systems describable by a single state variable,  $x$ . However, we now modify the system to include a disturbance term affecting the state dynamics, so that we may write

$$\begin{aligned} \dot{x} &= a(x) + b(x)u + g(x)d, \quad a(0) = 0 \\ z &= \begin{bmatrix} h(x) \\ u \end{bmatrix}, \quad h(0) = 0 \end{aligned} \quad (11.9)$$

where  $u$  is a scalar control,  $d$  is a scalar disturbance,  $z$  is a scalar penalized variable, and  $a$ ,  $b$ ,  $g$  and  $h$  are assumed to be analytic real-valued scalar functions of  $x$ . The control objective is to be accomplished by using the SDRE nonlinear  $H_\infty$  control technique:

- i. Write (11.9) in state-dependent coefficient (SDC) form

$$\begin{aligned} \dot{x} &= A(x)x + B(x)u + G(x)d \\ z &= \begin{bmatrix} H(x)x \\ u \end{bmatrix} \end{aligned} \quad (11.10)$$

- ii. Solve the nonlinear  $H_\infty$  SDRE

$$A(x)p(x) + p(x)A(x) + p(x)K(x)p(x) + H^2(x) = 0 \quad (11.11)$$

where  $K = (1/\gamma^2)G^2(x) - B^2(x)$ .

- iii. Construct the state feedback via

$$u = -B(x)p(x)x \quad (11.12)$$

Comparing the above with Chapter 5, we see that the only thing that has changed is that we now have  $K$  taking the place of  $-B^2$  in the scalar SDRE. We now investigate the results of this change in the required assumptions and nature of solutions for (11.11), and the corresponding local

stability analysis for the nominal unforced ( $d = 0$ ) closed loop system. In light of the discussion in Section 11.1, we expect the sign definiteness of  $K$  to figure prominently in the analysis. Recall that we assume analyticity of the system parameters  $a$ ,  $b$ ,  $g$ , and  $h$  and seek additional conditions under which the SDRE  $H_\infty$  control algorithm yields an *analytic* locally stabilizing state feedback. We consider the same four cases as in Chapter 5, proceeding analogously, except now we have  $d_K$  replacing  $2d_B$  as the smallest nonzero power of  $x$  multiplying the  $p^2$  term in (11.11), and  $c_K$  replacing  $-c_B^2$  as the associated nonzero coefficient. We note that for  $B$  and  $G$  assumed analytic,  $d_K$  will always be an even integer. In certain cases we must consider additional subcases corresponding to different assumed configurations of  $K$ . We shall denote these additional subcases per the following notation.

- i.  $d_B < d_G$ , so that  $d_K = 2d_B$  and  $c_K = -c_B^2$
- ii.  $d_B > d_G$ , so that  $d_K = 2d_G$  and  $c_K = (1/\gamma^2)c_G^2$
- iii.  $d_B = d_G$ , so that  $d_K = 2d_B = 2d_G$  and  $c_K = (1/\gamma^2)c_G^2 - c_B^2$

### 11.2.1 $H_\infty$ SDRE Solutions

Case 1 ( $2d_H < d_K$  and  $2d_H < d_A$ ) This case is unchanged from Chapter 5. The lowest-order nonzero part of (11.11) is

$$-c_H^2 x^{2d_H} = 0 \quad (11.13)$$

which, since  $c_H$  is nonzero, has no solution for all  $x$ . Obviously, there is again no stabilizing solution in this case, reinforcing that we cannot penalize powers of  $x$  smaller than those on which we may have some effect (either through the control or through the dynamics themselves).

Case 2 ( $d_A < 2d_B$ ) This case is also unchanged from Case 2 of Chapter 5. Since  $d_p \geq 0$  and  $d_A < 2d_B$ , the stability of the closed loop system is unaffected by the control, and is in fact determined by  $d_A$  and  $c_A$ . Invoking Lemma 5.2.1 we conclude that any solution that exists is stabilizing iff  $d_A$  is even and  $c_A < 0$ .

Case 3 ( $d_A = d_K$ )

Case 3A ( $2d_H > d_K$ ) In this case,  $d_p = 0$  and (11.11) reduces to

$$c_p(-c_K c_p - 2c_A)x^{d_A} + \dots = 0 \quad (11.14)$$

Thus, the two possible solutions to (11.14) are  $c_p = 0$ ,  $c_p = -2c_A/c_K$ . For  $c_p = 0$ , the leading-order control is  $u = 0$ , giving a closed loop system of  $\dot{x} = Ax$  (to 1st order). This solution is stabilizing iff the open loop system is stable ( $d_A$  is even and  $c_A < 0$ ). For  $c_p = -2c_A/c_K$ , the closed loop system becomes (to leading order)

$$\dot{x} = c_A x^{d_A+1} + c_B \left( -c_B \frac{-2c_A}{c_K} x^{d_A+1} \right) = \left( 1 + \frac{2c_B^2}{c_K} \right) c_A x^{d_A+1} \quad (11.15)$$

and we must consider the particular form of  $c_K$  dictated by the three possibilities given above to determine stability.

Case 3Ai This case is the same as for the nonlinear regulator. The closed loop system becomes

$$\dot{x} = -c_A x^{d_A+1} + \dots \quad (11.16)$$

which is stable iff  $d_A$  is even and  $c_A > 0$ . Thus, stabilizing solutions exist for this subcase iff  $d_A$  is even, which is always the case since  $d_A = d_K$ , and  $d_K$  is an even integer.

Case 3Aii With  $c_K = (1/\gamma^2)c_G^2$ , the closed loop system becomes

$$\dot{x} = \left( 1 + \frac{2\gamma^2 c_B^2}{c_G^2} \right) c_A x^{d_A+1} + \dots \quad (11.17)$$

which is stable iff  $d_A$  is even and  $c_A < 0$ . But if this is the case, the open loop system is also LAS. Thus, we have two stabilizing solutions in this case if  $c_A < 0$ , and no stabilizing solutions if  $c_A > 0$ . Note that we have  $K$  locally positive semidefinite for this subcase, and if  $c_A < 0$ , then  $\{A, B\}$  is stabilizable. From Theorem 11.1.3 we conclude that  $p = 0$  is the minimal, locally stabilizing solution to (11.11).

Case 3Aiii The closed loop system becomes

$$\dot{x} = \left( 1 + \frac{2c_B^2}{(1/\gamma^2)c_G^2 - c_B^2} \right) c_A x^{d_A+1} + \dots \quad (11.18)$$

which has the same solution and stability properties as Case 3Ai if  $c_K = (1/\gamma^2)c_G^2 - c_B^2 < 0$  and the same properties as Case 3Aii if  $c_K = (1/\gamma^2)c_G^2 - c_B^2 > 0$  (with different coefficients multiplying  $x^{d_A+1}$ , of course).

We now recall our discussion from Chapter 5 about the application of this subcase to the regulator problem when our cost function to be minimized is identically zero ( $d_H \rightarrow \infty$ ). For LTI open loop stable systems, the LQR optimal control for zero cost function is zero, and that is also what we see for SDRE nonlinear regulation and  $H_\infty$  control, if we take the minimal solution in the  $H_\infty$  setting. If an open loop LTI system is unstable and controllable, then the open loop unstable poles are moved to their stable mirror images in the left half complex plane (i.e., the real parts of the eigenvalues of the closed loop system are the negatives of their open loop counterparts). In the nonlinear regulator open loop unstable single state case, we saw that the closed loop dynamics were the negative of the open loop dynamics ( $a_{cl} = -a_{ol}$ ). In this SDRE nonlinear  $H_\infty$  case, we have a similar result, except now we have stabilization only when  $c_K = (1/\gamma^2)c_G^2 - c_B^2 < 0$ , and  $a_{cl} = -ka_{ol}$ , where  $k > 0$  is a perturbation factor reflecting the contribution of the  $(1/\gamma^2)c_G^2$  term in the Riccati solution.

**Case 3B** ( $2d_H = d_K$ ) In this case  $d_p$  again equals zero and the leading-order (positive) SDRE solution is given by

$$c_p = -c_A/c_K + \sqrt{c_A^2/c_K^2 - c_H^2/c_K} \quad (11.19)$$

Substituting (11.19) into  $u = -bpx$  and then into (11.9) we obtain the (leading-order) closed loop dynamics

$$\dot{x} = \left[ c_A - c_B^2 \left( -c_A/c_K + \sqrt{c_A^2/c_K^2 - c_H^2/c_K} \right) \right] x^{d_A+1} \quad (11.20)$$

We only consider the positive solution because it is trivial to show that the negative square root solution always yields an unstable closed loop system. We now again consider three subcases depending on the form of  $c_K$ .

**Case 3Bi** This is the same as the nonlinear regulator. We get the closed loop leading-order dynamics

$$\dot{x} = -\sqrt{c_A^2 + (c_B c_H)^2} x^{d_A+1} \quad (11.21)$$

which is always stable since  $d_A = d_K$  is even.

Case 3Bii The leading-order nonzero coefficient of  $p$  becomes

$$c_p = -\gamma^2 c_A / c_G^2 + \sqrt{\gamma^4 c_A^2 / c_G^4 - \gamma^2 c_H^2 / c_G^2} \quad (11.22)$$

so that we must have

$$\gamma^2 c_A^2 - c_H^2 c_G^2 > 0 \quad (11.23)$$

for a real solution to exist. If (11.23) is satisfied, we obtain the closed loop dynamics

$$\dot{x} = \left[ c_A + \gamma^2 c_B^2 / c_G^2 \left( c_A - \sqrt{c_A^2 - c_H^2 c_G^2 / \gamma^2} \right) \right] x^{d_A+1} + \dots \quad (11.24)$$

which is stable if  $c_A < 0$  and unstable if  $c_A > 0$ . Thus, a stabilizing solution can exist for this subcase iff the open loop dynamics are LAS, but the locally stabilizing solution is not  $p = 0$  as in Case 3A. Note that stronger state penalties and disturbance input effects (larger  $c_H^2$  and  $c_G^2$ ) make (11.23) harder to satisfy, but  $c_H^2$  is user-selectable and may be chosen to ensure existence of locally stabilizing solutions for some range of  $\gamma$ .

Case 3Biii This subcase is like subcase 3Aiii in that it is characterized by the sign of  $c_K = (1/\gamma^2)c_G^2 - c_B^2$ . If  $c_K < 0$ , then the situation is like Case 3Bi, and stabilizing solutions always exist, given by (11.19) and (11.20). If  $c_K > 0$ , then the situation is like Case 3Bii, in that a locally stabilizing, analytic solution exists iff

$$c_A^2 - c_H^2 c_K > 0 \quad (11.25)$$

and  $c_A < 0$ .

Case 4 ( $d_A > d_K$ )

Case 4A ( $2d_H = d_K$ ) In this case  $d_p = 0$  and (11.11) becomes

$$(c_H^2 + c_K c_p^2) x^{d_K} + \dots = 0 \quad (11.26)$$

giving

$$c_p^2 = -c_H^2 c_K \quad (11.27)$$

so that we must have  $c_K < 0$  for a real solution to exist. If this is the case, the closed loop dynamics become

$$\dot{x} = c_B^2 |c_H| / c_K x^{d_K+1} + \dots \quad (11.28)$$

which is always stabilizing since  $d_K$  is even and  $c_K < 0$ .

Case 4B ( $2d_H > d_K$ ) In this case  $d_p \geq 1$  and the existence of solutions is highly dependent on the particular structures of  $A$ ,  $K$ , and  $H$ . We thus give no criterion for the existence of stabilizing solutions in this case. We also note that such a case is somewhat unusual, in that we would be penalizing only large powers of  $x$  without penalizing the smaller, affectable powers.

### 11.2.2 Discussion

From the above analysis we see that, as predicted, the sign definiteness of  $c_K = (1/\gamma^2)c_G^2 - c_B^2$  plays a crucial role in determining when stabilizing solutions to the scalar  $H_\infty$  Riccati equation exist. When  $c_K < 0$ , the situation is analogous to the SDRE nonlinear regulator results of Chapter 5. This is true because when  $c_K < 0$ , controllability and stabilizability are equivalent for the  $K$  terms ((1,2)-blocks of  $\mathcal{H}$ ) of the regulator and  $H_\infty$  Riccati equations. Of course, observability and detectability are always equivalent for the two Riccati equations since they consider the same pair  $\{H, A\}$ . Thus, in the scalar case, if we have  $\{A(x), B(x)\}$  stabilizable and  $\{H(x), A(x)\}$  detectable near the origin, the nonlinear  $H_\infty$  control problem will be solvable for some (potentially large) value of  $\gamma$ , since we can guarantee  $c_K < 0$  by choosing  $\gamma$  large enough. This will not be true in general when we consider the multistate case, since directionality issues, and not just magnitude issues, will play a role in determining sign definiteness of  $K$ . We also saw that if  $c_K > 0$ , in some situations we were able to find LAS solutions, as predicted by Theorem 11.1.3. It remains to be seen whether these additional conditions can be uniformly identified and exploited. The results of this section again verify the known result of analytic and stabilizable/detectable systems yielding analytic stabilizing solutions to (11.11), but also show, just as in the regulator case, that stabilizability/detectability for

all  $x$  are not necessary in the scalar case, and illustrate under what conditions such solutions may or may not be obtained.

### 11.2.3 Examples

We now present three examples to illustrate the above theory, and to demonstrate the necessary and sufficient conditions for analytic stabilizing solutions derived herein.

#### Example 1

$$\dot{x} = x^3 + 5xu - xd; \quad h = cx^2 \quad (11.29)$$

This example is not controllable/observable, nor stabilizable/detectable in a neighborhood of the origin. Here we have  $A = x^2$ ,  $B = 5x$ ,  $G = -x$ , and  $H = cx$  so that  $d_A = 2$ ,  $c_A = 1$ ,  $d_B = 1$ ,  $c_B = 5$ ,  $d_G = 1$ ,  $c_G = -1$ ,  $d_H = 1$ , and  $c_H = c$  and Case 3Biii applies. Since  $d_A = 2$  is even, we expect stability of the closed loop to depend on the sign of  $c_K = (1/\gamma^2) - 25$ . For  $\gamma^2 > 1/25$ ,  $c_K < 0$  and we expect stable closed loop solutions. Solving (11.19) under this assumption (letting  $c_k = -k$ ,  $k > 0$ ) we find

$$c_p = 1/k + \sqrt{1/k^2 + c^2/k} \quad (11.30)$$

The resulting closed loop system is

$$\dot{x} = [1 - 25(1/k + \sqrt{1/k^2 + c^2/k})]x^3 + \dots \quad (11.31)$$

which is indeed (locally) stable since  $k < 25$ . If  $c_k > 0$  we need  $c_A < 0$  and  $c_A^2 - c_H^2 c_K = 1 - c^2 c_K > 0$  for existence of an LAS solution. Since  $c_A = 1 > 0$ , no stabilizing solutions exist for  $\gamma^2$  values smaller than  $1/25$ .

#### Example 2

$$\dot{x} = -x^3 + 5xu - xd; \quad h = cx^2 \quad (11.32)$$

All we have done here is to negate  $A$  from Example 1, so that now  $c_A = -1 < 0$ . Case 3Biii still applies. For  $\gamma^2 > 1/25$ ,  $c_K < 0$  and we expect stable closed loop solutions. Solving (11.19) as in

Example 1 we find

$$c_p = -1/k + \sqrt{1/k^2 + c^2/k} \quad (11.33)$$

and the resulting closed loop system

$$\dot{x} = [-1 - 25(-1/k + \sqrt{1/k^2 + c^2/k})]x^3 + \dots \quad (11.34)$$

which is indeed LAS since  $c_p = -1/k + \sqrt{1/k^2 + c^2/k} > 0$ . Now, for  $c_K > 0$ , since  $c_A < 0$ , we expect LAS solutions if  $1 - c^2 c_K > 0$ , or equivalently, if  $c_K = 1/\gamma^2 - 25 < 1/c^2$ . Thus, in this case we can achieve LAS closed loop systems with slightly better disturbance attenuation properties than in Example 1, because now we get LAS solutions for  $\gamma$  such that  $\frac{c^2}{25c^2+1} < \gamma < \frac{1}{25}$ .

### Example 3

$$\dot{x} = -x^2 + 5xu - xd; \quad h = cx^2 \quad (11.35)$$

Notice that all we have changed from Example 2 is  $a$  from  $-x^3$  to  $-x^2$  so that  $d_A = 1$  and  $d_K = 2d_B = 2$  so that  $d_A < 2d_B$  and Case 2 applies. Since  $d_A = 1$  is odd, we expect an unstable closed loop solution. Solving (11.11) we find  $d_p = 1$  and  $c_p = c^2/2$ , so that  $p$  is not even locally positive semidefinite. The resulting low-order control is

$$u = -5x\left(\frac{c^2}{2}x\right)x = -\frac{5c^2}{2}x^3 \quad (11.36)$$

giving the closed loop system

$$\dot{x} = -x^2 - \frac{5c^2}{2}x^4 + \dots \quad (11.37)$$

which as expected has stability properties unaffected by the control, and is clearly unstable.

From these examples we see the validity of the theory, and the important role  $c_K$  and  $d_K$  now play in local closed loop stability analysis. Indeed, this role of  $K$  as opposed to  $B$  in the SDRE is the main difference we observe between the regulator and the  $H_\infty$  theory, whereas the fact remains in both methods that we cannot penalize powers of  $x$  smaller than we can effect, nor can we stabilize systems with only terms in  $B(x)$  of order higher than  $A(x)$ .

### 11.3 Lyapunov Stability for Systems with Full Rank, Constant $B$ Matrices

In this section we are able to show that the global asymptotic stability properties of SDRE nonlinear regulators for systems with full rank, constant  $B$  matrices carries over to the SDRE nonlinear  $H_\infty$  unforced ( $d = 0$ ) closed loop system, with only two additional assumptions. The method of proof is the same as in Chapter 7, and is only slightly more complicated. Recall we consider the system

$$\begin{aligned} \dot{x} &= a(x) + Bu + Gd, \quad a(0) = 0 \\ z &= \begin{bmatrix} h(x) \\ \bar{R}u \end{bmatrix}, \quad h(0) = 0 \end{aligned} \quad (11.38)$$

where  $x, u \in \mathcal{R}^n, d \in \mathcal{R}^r, z \in \mathcal{R}^{2n}$ , and  $\bar{R}$  and  $B$  are nonsingular matrices. The vector functions  $a$  and  $h$  are assumed to be  $C^1$  real-valued functions of  $x$ , so that well-defined global SDC parametrizations are guaranteed to exist. The first assumption we make is implied in (11.38), namely that the disturbances enter the dynamics through a constant matrix  $G$ . The second assumption we make is that  $\gamma$  is selected in the suboptimal  $H_\infty$  control problem large enough so that

$$K = (1/\gamma^2)GG^T - BR^{-1}B^T < 0 \quad (11.39)$$

i.e., the (1,2)-block of the  $H_\infty$  Hamiltonian is negative definite, where  $R = \bar{R}^T \bar{R} > 0$ . Note that a finite value of  $\gamma$  satisfying (11.39) is guaranteed to exist by the positive definiteness of  $BR^{-1}B^T$ . We now formally state and prove the result.

**Theorem 11.3.1** *Consider the system (11.38) with  $a(x)$  and  $h(x)$  assumed to be  $C^1$  functions and  $G$  a constant matrix. Assume further that  $a(x) = A(x)x = 0 \Rightarrow x = 0$ ,  $\text{rank}(B) = n$  and  $H^T(x)H(x) > 0 \forall x$  where  $h(x) = H(x)x$ . Also, assume  $\bar{R}$  in (11.38) is constant and nonsingular, and (11.39) holds. Then application of the SDRE nonlinear  $H_\infty$  control algorithm defined by*

$$u = -R^{-1}B^T P(x)x \quad (11.40)$$

where  $P(x)$  is the stabilizing solution to

$$A^T(x)P(x) + P(x)A(x) + P(x)KP(x) + Q(x) = 0 \quad (11.41)$$

and  $K$  is as in (11.39), to (11.38) yields a closed loop system which is globally asymptotically stable.

**Proof:** Since (11.39) holds and  $H(x)$  is globally full rank, we note first of all that stabilizing solutions to (11.41) are guaranteed to exist for all  $x$  by Theorem 11.1.2, so that the control algorithm is globally well-defined. The negative definiteness of  $K$  also lets us write

$$K = -DD^T \quad (11.42)$$

for some nonsingular matrix  $D$ . With this definition (11.41) becomes

$$A^T(x)P(x) + P(x)A(x) - P(x)DD^T P(x) + Q(x) = 0 \quad (11.43)$$

Consider the globally positive definite, decrescent, and radially unbound Lyapunov function  $V = (1/2)q^T q$  where

$$q = D^{-1}x \quad (11.44)$$

Then we have  $\dot{V} = q^T \dot{q}$ . Differentiating (11.44) and using (11.40) we find

$$\dot{q} = D^{-1}\dot{x} = D^{-1}A(x)Dq - D^{-1}BR^{-1}B^T P(x)Dq \quad (11.45)$$

Similar to Chapter 7, define

$$\begin{aligned} A(q) &= D^{-1}A(x)D \\ P(q) &= D^T P(x)D \\ Q(q) &= D^T Q(x)D \end{aligned} \quad (11.46)$$

Using (11.46) and (11.45),  $\dot{V}$  becomes

$$\dot{V} = q^T [A(q) - D^{-1}BR^{-1}B^T D^{-T} P(q)]q \quad (11.47)$$

so that

$$\dot{V} \leq (\mu_2[A(q)] - \lambda[D^{-1}BR^{-1}B^T D^{-T} P(q)])q^T q \quad (11.48)$$

where  $\mu_2(A) = \max[\operatorname{Re}\lambda(A+A^T)/2]$  is the matrix measure of  $A$  with respect to the Euclidean norm. We note that (11.48) is different from the corresponding expression for  $\dot{V}$  in the regulator case only by the presence of the  $D^{-1}BR^{-1}B^TD^{-T}$  premultiplying  $P(q)$ . Now, using (11.46), (11.43) becomes

$$A^T(q)P(q) + P(q)A(q) - P(q)P(q) + Q(q) = 0 \quad (11.49)$$

so that using Mori's lower bound [49] for  $\underline{\lambda}[P(q)]$ , we find  $\underline{\lambda}[P(q)] > \mu_2[A(q)]$  or equivalently  $\mu_2[A(q)] - \underline{\lambda}[P(q)] < 0$ , just as in the regulator case, where we recall we may use Mori's bound because  $A(q)$  is guaranteed nonsingular by assumed global nonsingularity of  $A(x)$  and (11.46). Now, since  $\underline{\sigma}(WX) \geq \underline{\sigma}(W)\underline{\sigma}(X)$  for any two matrices  $W$  and  $X$  [57], then  $-\underline{\lambda}[D^{-1}BR^{-1}B^TD^{-T}P(q)] \leq -\underline{\lambda}[D^{-1}BR^{-1}B^TD^{-T}]\underline{\lambda}[P(q)]$ . Thus, if  $\underline{\lambda}[D^{-1}BR^{-1}B^TD^{-T}] \geq 1$ , then by (11.48) we will have  $\dot{V} < 0$  for all  $q$ , and the theorem will be proven. Now, recall that by definition  $K = (1/\gamma^2)GG^T - BR^{-1}B^T = -DD^T$ , so that we may write

$$BR^{-1}B^T = DD^T + (1/\gamma^2)GG^T$$

Thus, we have

$$\begin{aligned} \underline{\lambda}[D^{-1}BR^{-1}B^TD^{-T}] &= \underline{\lambda}[D^{-1}(DD^T + (1/\gamma^2)GG^T)D^{-T}] \\ &= \underline{\lambda}[I + (1/\gamma^2)D^{-1}GG^TD^{-T}] \\ &= 1 + (1/\gamma^2)\underline{\lambda}[D^{-1}G(D^{-1}G)^T] \geq 1 \end{aligned} \quad (11.50)$$

since the minimum eigenvalue of the positive semidefinite matrix  $D^{-1}G(D^{-1}G)^T$  is greater than or equal to zero, and the theorem is proven. ■

We conclude this section by noting that the other theorems in Chapter 7 may be similarly extended to the SDRE nonlinear  $H_\infty$  case by making the assumption (11.39), since those theorems also establish conditions under which  $\mu_2[A(q)] - \underline{\lambda}[P(q)] < 0$ , and (11.50) holds. Thus, all the results of Chapter 7 apply to the SDRE nonlinear  $H_\infty$  control problem, provided we have negative definiteness of the constant matrix  $K = (1/\gamma^2)GG^T - BR^{-1}B^T$ .

#### 11.4 Asymptotic Stability of Sampled Data SDRE Nonlinear $H_\infty$ Controllers

In this section we give the changes in assumptions necessary to extend the results of Chapters 8 and 9 to the closed loop system obtained by applying the sampled data SDRE nonlinear  $H_\infty$  control algorithm. We also outline some slight additional considerations needed to complete the proofs. Our control algorithm is determined analogously to the sampled data SDRE nonlinear regulator. That is, at sampling time  $t_k$  we solve the sampled data nonlinear  $H_\infty$  SDRE

$$A_k^T P_k + P_k A_k + P_k K_k P_k + Q_k = 0 \quad (11.51)$$

where  $A_k = A(x(t_k))$  and likewise for all other variables,  $Q = H^T H$ , and  $K(x) = (1/\gamma^2)G(x)G^T(x) - B(x)R^{-1}(x)B^T(x)$  as in the previous sections of this chapter, and apply the constant control

$$u_k = -R_k^{-1}B_k^T P_k x_k \quad (11.52)$$

until the next sampling time  $t_{k+1}$ . With  $H$  selected globally nonsingular, nonlinear  $H_\infty$  versions of Theorems 8.3.1 and 8.4.1 apply provided the following three additional assumptions are made.

- i.  $G(x)$  is globally analytic with respect to  $x$
- ii. The pair  $\{A(x), K(x)\}$  is globally stabilizable
- iii. Stabilizing solutions to (11.51) exist everywhere along the trajectory

Assumption i is a new assumption, which is sufficient to ensure  $K(x)$  is globally analytic with respect to  $x$ . Assumption ii is in addition to the previous assumption of global stabilizability of  $\{A(x), B(x)\}$ , which is no longer sufficient to guarantee that solutions of the  $H_\infty$  Riccati equation will be globally analytic with respect to  $x$ . Thus, Assumptions i and ii, along with the other assumptions in Theorem 8.3.1, combine to ensure the global analyticity of  $P$ , which is crucial to the proofs of both Theorems 8.3.1 and 8.4.1. At this point we remark that Assumption i is no more restrictive than the assumption that the other system parameters are globally analytic, but Assumption ii could be potentially difficult to satisfy, particularly if the disturbance inputs enter the system in directions which are not controllable. If, on the other hand, at each  $x$  we have  $G(x)$  is

contained in  $\text{span}(B(x))$ , then stabilizability of  $\{A, B\}$  guarantees stabilizability of  $\{A, K\}$  for some range of  $\gamma$  values. As mentioned throughout this chapter, we must be concerned with existence of solutions to (11.51), since stabilizability of  $\{A, B\}$  and  $\{A, K\}$  and detectability of  $\{H, A\}$ , which we assume hold, are not by themselves enough to guarantee that stabilizing solutions to (11.51) exist. Thus, we need Assumption iii to guarantee well-posedness of the control algorithm along the trajectory. Of course, in view of Assumption ii and Theorem 11.1.2, Assumption iii can be satisfied by assuming that  $K \leq 0$  everywhere along the trajectory, although this can be a quite restrictive assumption. With these modified assumptions the proof of Theorem 8.3.1 remains unchanged, and the proof of Theorem 8.4.1 has only a slight modification in the computation of  $\Delta V_k$ , which is dictated by the change in form of the Riccati equation. Instead of

$$\begin{aligned}\Delta V_k &\leq s_{k+1} \delta_k x_k^T \left[ \epsilon I + F_k^T P_k + P_k F_k \right] x_k + x_k^T \mathcal{O}(\delta^2) x_k \\ &= s_{k+1} \delta_k x_k^T \left[ \epsilon I - Q_k - P_k B_k R_k^{-1} B_k^T P_k \right] x_k + x_k^T \mathcal{O}(\delta^2) x_k\end{aligned}\quad (11.53)$$

as in the SDRE nonlinear regulator, (11.51) rearranges to give

$$F_k^T P_k + P_k F_k = -Q_k - P_k B_k R_k^{-1} B_k^T P_k - (1/\gamma^2) P_k G_k G_k^T P_k \quad (11.54)$$

which is still negative definite since  $Q_k < 0$ . Thus, we get

$$\Delta V_k \leq s_{k+1} \delta_k x_k^T \left[ \epsilon I - Q_k - P_k B_k R_k^{-1} B_k^T P_k - (1/\gamma^2) P_k G_k G_k^T P_k \right] x_k + x_k^T \mathcal{O}(\delta^2) x_k \quad (11.55)$$

for the SDRE nonlinear  $H_\infty$  case, so that the arbitrary tolerance  $\epsilon$  being selected to be less than  $\lambda[Q_k]$  is still sufficient to guarantee negativity of (11.55). Indeed, the extra term  $-(1/\gamma^2) P_k G_k G_k^T P_k$  in (11.55) can only help to make  $\Delta V_k$  more negative.

If we now consider the results of Chapter 9, we see that they, too, remain unchanged, provided the above three additional assumptions are made. We could, if we so desired, redefine the Lyapunov function scaling factor  $s_r$  for  $t_r \in (t_k, t_{k+1}]$  according to

$$s_r = \frac{x_k^T P_k x_k}{x_k^T P_r x_k} s_k \quad (\text{if } x_k^T P_k' x_k \geq x_k^T [Q_k + P_k B_k R_k^{-1} B_k^T P_k + (1/\gamma^2) P_k G_k G_k^T P_k] x_k) \quad (11.56)$$

$$= s_k \quad (\text{otherwise}) \quad (11.57)$$

where  $P'_k$  is as before and  $s_0 = 1$ , to take advantage of the additional negative contribution of the  $-(1/\gamma^2)P_k G_k G_k^T P_k$  term to (11.55), but there really is no need to do so. If we leave  $s_r$  defined as in Chapter 9, then although the definition of  $E$  would change to

$$E = \{x \in \mathcal{R}^n \mid x^T [Q + PBR^{-1}B^T P + (1/\gamma^2)PGG^T P]x = 0 \text{ or } s(x) = 0\} \quad (11.58)$$

$E$  remains the same, since  $P(x)x = 0$  causes both the second and third terms in the first expression in (11.58) to vanish, so that the theorems of Chapter 9 would still apply to the  $d = 0$  nonlinear  $H_\infty$  case, provided the three additional assumptions above are made. We conclude this section by noting that, for the results of Section 9.3.6 to be valid in the nonlinear  $H_\infty$  setting, we must have both  $\{A, B\}$  and  $\{A, K\}$  globally stabilizable, and that is why we stated above that Assumption ii was in addition to, and not instead of, having  $\{A, B\}$  globally stabilizable.

### 11.5 Exponential Stability of Nonlinear SDRE $H_\infty$ Systems

Since the arguments of Chapter 10 rely on stability and analyticity of the matrix function  $F = A + BR^{-1}B^T P$ , the assumptions of Section 11.4, which guarantee those properties for the  $d = 0$  SDRE nonlinear  $H_\infty$   $F$  matrix function, are also sufficient to extend those exponential stability arguments to the  $H_\infty$  case. All other assumptions made in the case of the nonlinear regulator, such as diagonalizability of  $F$ , would also need to hold, of course. The only other consideration that is different in the  $H_\infty$  case is that the use of lower bounds on  $\underline{\lambda}[P]$  to make the exponential decay factor  $m_r$  larger by increasing  $\underline{\lambda}[Q]$  require the additional assumption that  $K = (1/\gamma^2)GG^T - BR^{-1}B^T \leq 0$ . We note that this assumption might already be made in order to satisfy Assumption iii of Section 11.4, and thus may represent no additional requirement.

In this chapter we have shown that the theory previously developed for the SDRE nonlinear regulator may be extended to the nonlinear  $H_\infty$  setting, provided suitable additional assumptions on the system parameters are made. Although some of these assumptions can be quite restrictive, they are no more so than in the corresponding linear case, and at least give a sufficient set of conditions which guarantee well-posedness of the algorithm.

## *XII. Design Problem*

In this chapter we examine an Air-Force-relevant nonlinear control problem of nontrivial state dimension that we wish to solve by the methods of Chapter 2. The chosen problem involves momentum control of an axial gyrostat, with various assumed configurations. As discussed in Section 1.2, the chosen problem exhibits highly nonlinear dynamics and limited controllability. It is therefore a good test of the applicability of modern nonlinear control design techniques to somewhat realistic problems.

### *12.1 Problem Description*

One important satellite attitude stabilization technique is dual-spin stabilization. A dual-spin satellite consists of two bodies capable of relative rotation, with one body spinning relatively fast (the rotor) to provide stabilization, and one body (the platform) spinning relatively slowly in order to perform mission requirements (i.e., to remain earth-pointing). Typical deployment scenarios result in both bodies initially spinning at nearly the same rate about a single axis (the so-called all-spun condition), so that some type of spinup maneuver is required to despin the platform. This spinup maneuver is described by strongly nonlinear equations of motion, which are complicated by unbalance or asymmetry of either the platform or the rotor, the existence of which leads to several interesting and often undesirable phenomena. Spinup dynamics of satellites under these conditions of imbalance or asymmetry have received significant attention in the literature, and the reader is referred to [23] for a complete survey. Two particular cases of interest are commonly studied, corresponding to assumptions of both bodies being dynamically balanced, but with one body axisymmetric and the other asymmetric. These spinup maneuvers are typically performed by applying a small, constant internal torque to the platform to decrease its rate of rotation, thereby transferring angular momentum to the rotor. Under this small, constant internal torque assumption, one particularly undesirable phenomenon observed for the case of axisymmetric platform and asymmetric rotor is

called *resonance capture* or sometimes *precession phase lock*. Resonance capture happens when a dual-spin spacecraft that starts in the all-spun condition (with nearly zero cone or nutation angle) departs from that condition upon execution of the despin maneuver, creating a large cone angle. Hall [22] has derived moment of inertia related conditions for when resonance capture can occur, and has shown that the phenomenon is dependent on initial conditions. An interesting nonlinear control problem is therefore to try to accomplish the despin maneuver by use of other than a constant control torque, avoiding the phenomenon of resonance capture in the process, or at least extending the neighborhood of initial conditions for which resonance capture is avoided. In the dual case of asymmetric platform and axisymmetric rotor, the small constant torque is used to spin up the rotor, thereby reducing the angular momentum and hence angular velocity of the platform. For this set of assumptions we obtain subcases corresponding to oblate, prolate, and transverse spinup maneuvers, each of which has distinguishing interesting characteristics. In particular, oblate and prolate spinup maneuvers are associated with satellites possessing radically different inertia characteristics (defined in Section 12.2, yet both start from relatively small cone angles ( $< 25$  degrees). Transverse spin up, on the other hand, represents a large cone angle regulation problem which can be attempted for either oblate or prolate spacecraft. Hall and Rand [25] have observed differing adverse behaviors for these three maneuvers under the small, constant torque assumption, which are associated with the number of open loop separatrix crossings each maneuver encounters, where by separatrix we mean a manifold which separates domains of action of different equilibria. The oblate, prolate, and transverse spinups encounter zero, one, and two such crossings, respectively. Oblate spinup has no separatrix crossing and thus no adverse effect on the final cone angle is seen. The prolate and transverse cases, however, have crossings which generally result in larger final cone angles, and the phenomenon is typically worse for the prolate case under the small torque assumption. In this chapter we attack this second set of problems, attempting the three types of spinup maneuvers described by the technique of SDRE nonlinear regulation. For comparison purposes, we also demonstrate

how SDRE nonlinear  $H_\infty$  control, feedback linearization, and recursive backstepping theory may be applied to this problem.

## 12.2 Equations of Motion

A simple model of dual-spin spacecraft which we first propose to use is the axial gyrostat. An axial gyrostat is a coupled rigid body system, in which the relative rotation between bodies is constrained to occur about an axis of symmetry for at least one of the bodies. In accordance with the above discussion and to allow direct comparison of our results with the analysis of [25], we make the following three assumptions:

- i. Both the rotor and platform are dynamically balanced.
- ii. The rotor is axisymmetric and its relative spin axis is parallel to a principal axis of the gyrostat.
- iii. The platform is asymmetric.

The simplified satellite dynamics for an axial gyrostat under the above assumptions consist of a fourth-order nonlinear dynamic single-input system, and are given in [25]. Before specifying them here, we need some definitions. We define a principal body axis coordinate frame located at the center of mass of the gyrostat by the vectors  $e_1, e_2, e_3$  relative to inertial space, where  $e_1$  is the axis about which the platform and rotor may have relative rotation (the spin axis), and  $e_2$  and  $e_3$  are principal axes orthogonal to  $e_1$  (see Figure 12.1). We define the scalar  $\mu$  as the axial angular momentum of the rotor (about the  $e_1$  axis), and the vector  $x$ , containing three elements, as the total angular momentum of the gyrostat about the  $e_i$  axes. An interesting system characteristic is that, since there are no external torques acting on the gyrostat, we have conservation of angular momentum. Thus,  $\|x\|$  is a constant, and in fact we make  $x$  dimensionless by scaling so that  $\|x\| = 1$ . The control input  $u$  will be the scalar torque applied to the rotor. We also define dimensionless inertia parameters  $i_j$  via

$$i_j = 1 - \frac{I_p}{I_j}, \quad j = 1, 2, 3 \quad (12.1)$$

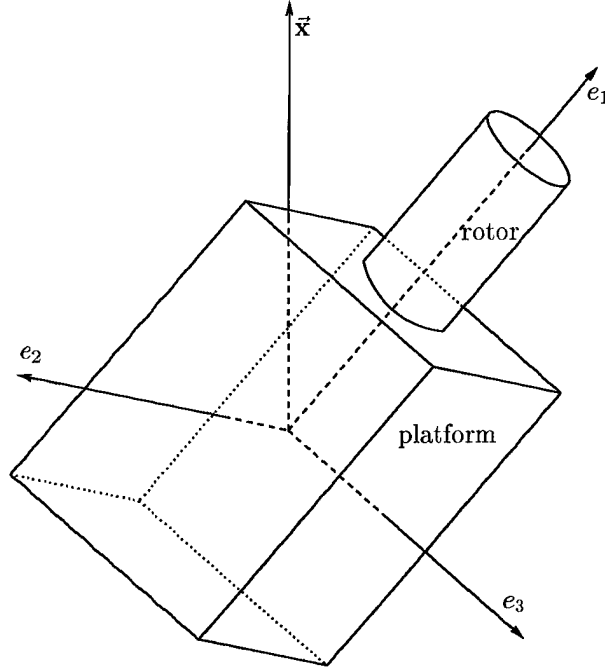


Figure 12.1: Gyrostat Model of Axial Dual-Spin Spacecraft

where  $I_p = I_1 - I_s$ ,  $I_s$  is the axial moment of inertia of the rotor, and  $I_j$  are the principal moments of inertia of the gyrostat. Using these definitions, the system dynamics for the gyrostat are

$$\begin{aligned}
 \dot{x}_1 &= (i_2 - i_3)x_2x_3 \\
 \dot{x}_2 &= (i_3x_1 - \mu)x_3 \\
 \dot{x}_3 &= -(i_2x_1 - \mu)x_2 \\
 \dot{\mu} &= u
 \end{aligned} \tag{12.2}$$

which, interestingly, are invariant under the change of variables

$$(x_1, x_2, x_3, i_2, i_3) \rightarrow (-x_1, x_3, -x_2, -i_3, -i_2) \tag{12.3}$$

We note that  $i_1$  does not appear in the system equations, but that it does impact the system via initial conditions for spinup. This can be seen from the fact that in the all-spun condition, the initial conditions are  $(x_1, x_2, x_3, \mu) \approx (1, 0, 0, i_1)$ . We also see from (12.1) that  $i_1 = I_s/I_1$ , so that  $i_1$  is always positive. The relationship between the other two inertia parameters determines whether the spacecraft is oblate or prolate. For oblate spacecraft we have  $i_3 < i_2 < 0$  or equivalently

$I_p > I_2 > I_3 > 0$ , whereas for prolate spacecraft we have  $i_2 > i_3 > 0$  or equivalently  $I_2 > I_3 > I_p$ . We note that due to the freedom in choosing  $e_2$  and  $e_3$ , the above assumption  $i_2 > i_3$  represents no loss of generality. Finally, the cone or nutation angle  $\eta$  (the angle between the  $e_1$  axis and the angular momentum vector,  $x$ ) is given by the relationship

$$\cos \eta = x_1 \Rightarrow \eta = \arccos x_1 \quad (12.4)$$

### 12.3 Design Objectives

From a controls perspective, we have three desired maneuvers to perform, each of which is characterized by inertia parameters, initial states, and desired final states. In oblate spinup we start with  $x_1 \approx 1$  and  $\mu = i_1 x_1$ , so that we have a small initial cone angle, and we desire to drive (12.2) to the final state  $(x_1, \mu) = (1, 1)$ , since by (12.4) and the relationship

$$x_1 = I_p \omega_1 + \mu \quad (12.5)$$

this results in a despun platform with zero cone angle ( $\omega_1$  is the angular velocity of the platform about the  $e_1$  axis). As mentioned above, the system has been scaled so that  $\|x\| = 1$ . This implies that if we achieve the above objective, we must have  $x_2 = x_3 = 0$ , so that actually we desire to drive the system to  $(x_1, x_2, x_3, \mu) = (1, 0, 0, 1)$ . In prolate spinup, we again start with  $x_1 \approx 1$ ,  $\mu = i_1 x_1$ , and desire to drive the system to the final state  $(x_1, x_2, x_3, \mu) = (1, 0, 0, 1)$ , but we have different inertia parameters. In the transverse spinup maneuver, we start with  $x_1 \approx 0$ ,  $\mu = i_1 x_1$ , and we again desire to drive the state to  $(x_1, x_2, x_3, \mu) = (1, 0, 0, 1)$ . The difference between this maneuver and the oblate and prolate spinup maneuvers is thus the much larger initial cone angle in the transverse spinup maneuver. This maneuver represents recovery from a flat spin, for example, and should most seriously stress the control strategies.

We note that these control problems do not qualify as normal unconstrained regulation problems, but are instead constrained nonzero setpoint problems. It is expected that the constraint will manifest itself as a lack of controllability in the  $x$  states, which may cause problems for some of the

methods of Chapter 2. Various approaches may therefore be considered to attempt to circumvent this problem, some of which we mention here.

By implicitly recognizing the constraint  $\|x\| = 1$  on the control problem, we might consider attempting only the regulation of  $(x_2, x_3)$  to  $(0, 0)$ , using the constraint to thereby drive  $x_1$  to one. In the nonlinear regulator setup, for example, this would require not penalizing deviations of  $x_1$  from zero, so that detectability of  $x_1$  becomes necessary for existence of a stabilizing solution to the Riccati equation. Another complication of this approach is that even if  $x_2$  and  $x_3$  are successfully regulated, under the constraint  $x_1$  remains free to take on either of the values plus or minus one.

Another conceptual approach might be to change coordinates by defining  $\hat{x}_1 = x_1 - 1$ , so that in the new coordinates we do indeed have a true regulation problem. However, this option changes the nature of the constraint from  $\|x\| = x_1^2 + x_2^2 + x_3^2 = 1$  to  $\hat{x}_1^2 + 2\hat{x}_1 + x_2^2 + x_3^2 = 0$ . Successful regulation of  $\hat{x}_1$  still drives both  $x_2$  and  $x_3$  to zero, however, as desired.

A third option would be to use the constraint to eliminate one state variable from the dynamics. However, since the constraint only involves squares of the state variables, this option would require taking square roots, the appropriate signs of which might be difficult to determine.

In all of the above approaches, we still have left unaddressed the nonzero setpoint problem of driving  $\mu$  to one. This problem is easily solved by a change of coordinates defined by letting  $\nu = \mu - 1$ . This part of the control problem then reduces to regulation of  $\nu$ , which fits readily into the SDRE nonlinear regulator framework.

## 12.4 Open Loop System Analysis

Before developing the control design, we study the open loop model (12.2) and its relevant properties. From Chapter 9, we know that open loop equilibrium points can be important to closed loop stability analysis. We therefore observe that, for  $i_2 \neq i_3$ , the only open loop equilibrium points that are independent of  $\mu$  are  $x = (\pm 1, 0, 0)$ . Other equilibria exist for fixed values of  $\mu$ , but since  $\mu$  is

obviously completely controllable and we have a fixed desired final value for it, we are interested only in the  $\mu$ -independent equilibria specified above.

Now, from Section 4.7, we know that the system must be nonlinearly stabilizable for any control algorithm to work. We thus deepen our analysis by studying the nonlinear controllability of (12.2). Performing the iterative procedure for determining the Control Lie Algebra described in Chapter 6, it is easily found (using Mathematica [73] for instance) that

$$\Delta_3 = \begin{bmatrix} 0 & 0 & x_2^2 - x_3^2 & 2x_2x_3(2\mu - i_2x_1 - i_3x_1) \\ 0 & -x_3 & -x_1x_2 & x_3(-2\mu x_1 + 2i_3x_1^2 + i_2x_2^2 - i_3x_3^2) \\ 0 & x_2 & x_1x_3 & x_2(-2\mu x_1 + 2i_2x_1^2 - i_2x_2^2 + i_3x_3^2) \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (12.6)$$

so that on an open and dense subset of  $\mathcal{R}^4$  the gradients of invariant, uncontrollable coordinates of (12.2) are given by

$$\Delta_3^\perp = \{y \in \mathcal{R}^4 \mid y^T \Delta_3 = 0\} \quad (12.7)$$

From (12.6) first observe that  $\mu$  is always controllable. Thus, uncontrollable states will always lie in the  $(x_1, x_2, x_3)$  space. Also, it is trivially verified that for all  $y \in \mathcal{R}^4$ ,  $y = [x_1 \ x_2 \ x_3 \ 0]^T$  belongs to  $\Delta_3^\perp$ . Thus,  $\Delta_3$  is rank three for all  $x$ , and there exists a coordinate in the state space  $\phi(x)$  such that  $d\phi = y$ , and  $\dot{\phi}$  is unaffected by  $u$ . Solving for  $\phi$  we find

$$\phi = c(x_1^2 + x_2^2 + x_3^2) \quad (12.8)$$

so that we have recovered the constraint due to conservation of angular momentum, which is unchangeable by any choice of control, so that in fact  $\dot{\phi} = 0$ . What this means is that all trajectories must remain on the unit momentum sphere centered at  $(x_1, x_2, x_3) = (0, 0, 0)$  (plus the controllable  $\mu$  component trajectory). This uncontrollable coordinate, by itself, does not prevent us from reaching the desired equilibrium state since it, of course, lies on the sphere, and thus, does not imply that the system is not nonlinearly stabilizable.

Continuing our examination of (12.6), we see that if  $x_2 = x_3 = 0$ , then the uncontrollable space has dimension three, and, in fact, consists of the  $(x_1, x_2, x_3)$  space. What this means is that once we

hit either of the equilibrium surfaces  $(\pm 1, 0, 0, \mu)$ , we are stuck there forever, regardless of how we select the control. The implication here is that, if we seek to drive the system to  $x_1 = 1$ , we must avoid trajectories passing through the equilibrium at  $x_1 = -1$ , and vice versa.

The above two cases are clearly deduced from examining the determinant of  $\Delta_3$ , which is given by

$$\det[\Delta_3] = x_2 x_3 (x_2^2 - x_3^2)(0) \quad (12.9)$$

The zero in (12.9) comes from the global one-dimensional rank deficiency of (12.6), which gives a globally one-dimensional uncontrollable space. For  $x_2 = x_3 = 0$ , the loss of rank in (12.6) is three, again giving a correspondingly dimensioned uncontrollable space. From (12.9) it can be seen that, for other locations in the state space (namely  $x_2 = 0$  or  $x_3 = 0$ , and  $x_2^2 = x_3^2$ ), that  $\Delta_3$  may also lose rank of degree more than one. However, there are no other possible invariant trajectories in the state space when  $\mu \rightarrow 1$  for which the dimension of  $\Delta_3$  remains constant, and this implies that no other invariant, uncontrollable manifolds exist near the desired equilibrium point. To illustrate this concept, let  $x_2 = 0$ ,  $x_3 \neq 0$ . Then  $\Delta_3$  becomes

$$\Delta_3 = \begin{bmatrix} 0 & 0 & -x_3^2 & 0 \\ 0 & -x_3 & 0 & x_3(-2\mu x_1 + 2i_3 x_1^2 - i_3 x_3^2) \\ 0 & 0 & x_1 x_3 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (12.10)$$

In this situation,  $\Delta_3^\perp = [x_1 \ 0 \ x_3 \ 0]$ . Thus, the uncontrollable space is the circle  $x_1^2 + x_3^2 = 1$ . However, to stay on this circle requires  $\dot{x}_2 = (i_3 x_1 - \mu)x_3 = 0$ . But by assumption  $x_3 \neq 0$  (or else we return to the case already analyzed above), so that we must have  $i_3 x_1 - \mu = 0$  or

$$\mu = i_3 x_1 \quad (12.11)$$

Now, by (12.1)  $i_3$  is strictly less than one, and so is  $x_1$  (for  $x_3 \neq 0$ ), so that as we drive the completely controllable  $\mu$  toward one we must eventually violate (12.11) and leave the  $x_2 = 0$  constraint. A similar analysis for the  $x_3 = 0$ ,  $x_2 \neq 0$  and  $x_2^2 - x_3^2 = 0$  cases hold. From this analysis we conclude that, although the system (12.2) is not globally nonlinearly stabilizable, we can expect reasonable

success in driving the system to the desired equilibrium point if our initial conditions are sufficiently far away from the undesired equilibrium point at the opposite pole of the momentum sphere.

We can obtain one characterization of 'sufficiently far' by algebraically removing one of the  $x$  states, and repeating the above analysis on the reduced system. To this end we define the coordinate transformation  $q_1 = x_3$ ,  $q_2 = x_2$ ,  $q_3 = \phi$ , with  $\phi$  defined as in (12.8) with  $c = 0.5$ , and eliminate  $x_1$  as an independent variable by writing

$$x_1(q) = \pm \sqrt{1 - q_1^2 - q_2^2} \quad (12.12)$$

Since the Jacobian of this mapping is

$$J = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ x_1 & x_2 & x_3 \end{bmatrix} \quad (12.13)$$

we see by the Inverse Function Theorem that the mapping is not one-to-one in a neighborhood of  $x_1 = 0$ . Thus, we need to know which hemisphere of the momentum sphere we are in to complete the mapping, so as to choose the appropriate sign in (12.12). With this coordinate change the equations of motion become

$$\begin{aligned} \dot{q}_1 &= (\mu - i_2 x_1) q_2 \\ \dot{q}_2 &= (i_3 x_1 - \mu) q_1 \\ \dot{\mu} &= u \end{aligned} \quad (12.14)$$

where  $x_1$  is defined as in (12.12), and we have eliminated the trivial state equation  $\dot{q}_3 = 0$ . The nonlinear controllability procedure for this system yields

$$\Delta_2 = \begin{bmatrix} 0 & q_2 & q_1 x_1 \\ 0 & -q_1 & q_2 x_1 \\ 1 & 0 & 0 \end{bmatrix} \quad (12.15)$$

which has determinant

$$\det[\Delta_2] = x_1(q_1^2 + q_2^2) \quad (12.16)$$

This distribution thus has full rank as long as  $x_1 \neq 0$  and  $q_1 = x_3$  and  $q_2 = x_2$  are not both zero. Summarizing, we see that if we restrict  $x$  to remain in a hemisphere of the momentum sphere (either  $x_1 > 0$  or  $x_1 < 0$ ), then our coordinate change is well-defined (we know what sign to take in (12.12)), and the only place we lose controllability is at the desired equilibrium  $(q_1, q_2) = (0, 0)$ .

The final issue we wish to explore in open loop analysis is the existence of stabilizable and detectable factorizations of (12.2), which we require in order to guarantee well-posedness of the control algorithm. Since (12.2) is globally analytic with respect to  $x$ , we know that in a neighborhood of the origin, the SDRE regulator reverts to the linear quadratic regulator acting on the linearization of (12.2). A necessary condition for well-posedness of problems for numerical SDRE methods, as mentioned in Section 2.4.3, is thus that the system to be controlled have a stabilizable and detectable linearization. We also know from Chapter 6 that, for any SDC dynamics parametrization  $a(x) = A(x)x$  of (12.2),  $A(0)$  must equal the Jacobian of  $a$  evaluated at zero (and likewise for  $h(x) = H(x)x$ ), so that this necessary condition becomes stabilizability of  $\{J(0), B(0)\}$  and detectability of  $\{H(0), J(0)\}$ . Computing the Jacobian of (12.2) we find

$$J(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (12.17)$$

so that the linearized controllability matrix is given by

$$M_{cf}(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (12.18)$$

The linearization is obviously not controllable, so that detectability requires left half plane eigenvalues for all three  $x$  states. From (12.17) we see that all three of these eigenvalues are zero, however. The SDRE control is thus not well-defined for this system in a neighborhood of the origin. Recall, however, that we will never actually approach the origin since all trajectories must lie on the unit

momentum sphere, and that actually we need to shift the origin to make the desired closed loop equilibrium point our origin, as discussed in Section 12.3. To do so we define

$$\hat{x}_1 = x_1 - 1 \quad (12.19)$$

$$\nu = \mu - 1 \quad (12.20)$$

transforming (12.2) into

$$\begin{aligned} \dot{\hat{x}}_1 &= (i_2 - i_3)x_2x_3 \\ \dot{x}_2 &= (i_3\hat{x}_1 - \nu)x_3 + (i_3 - 1)x_3 \\ \dot{x}_3 &= -(i_2\hat{x}_1 - \nu)x_2 - (i_2 - 1)x_2 \\ \dot{\nu} &= u \end{aligned} \quad (12.21)$$

which has Jacobian at zero

$$J(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & i_3 - 1 & 0 \\ 0 & 1 - i_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (12.22)$$

and controllability matrix the same as (12.18). Regardless of the values of  $i_2$  and  $i_3$ , we see that (12.22) always has a zero eigenvalue for  $x_1$ , corresponding to the constraint manifold  $x_1^2 + x_2^2 + x_3^2 = 1$ . Thus, for well-posedness of the SDRE algorithm, we need to remove the  $x_1$  state from the system equations, and return to a system of the form (12.14). However, we still need to incorporate (12.20), giving the transformed system equations

$$\begin{aligned} \dot{q}_1 &= (\nu - i_2x_1)q_2 + q_2 \\ \dot{q}_2 &= (i_3x_1 - \nu)q_1 - q_1 \\ \dot{\nu} &= u \end{aligned} \quad (12.23)$$

which has Jacobian at zero

$$J(0) = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (12.24)$$

and linearization controllability matrix

$$M_{cf}(0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (12.25)$$

From (12.24) we see that still the linearization has eigenvalues on the imaginary axis, so that the requirements of stabilizability are not satisfied near the origin, where from now on the origin is the desired closed loop equilibrium in the transformed state space. As one might expect from the presence of imaginary axis eigenvalues in (12.24), application of the SDRE nonlinear regulation algorithm to (12.23) with small initial cone angle and only the  $\nu$  state penalized results in  $\nu$  being driven to zero, while the  $q_1$  and  $q_2$  states enter and remain in an oscillatory limit cycle about the origin. If we penalize all three states, it is interesting that all three states end up in a limit cycle. Typical simulation plots for these cases are given in Section 12.6.

**Remark:** *Unfortunately, this is the best we can do with the system as is, which points out an interesting potential limitation of the numerical SDRE methods. This potential limitation is that, for some systems with linearly uncontrollable purely imaginary open loop eigenvalues in their linearizations, it may be possible to stabilize the corresponding modes with a nonlinear control law. The numerical SDRE methods, however, revert to standard linear design methods near the origin, and so will not in general stabilize such systems. We conjecture that this limitation could possibly be overcome by modifying the SDRE algorithm to allow neutrally stabilizing solutions [56] to AREs (the closed loop  $F$  matrix has eigenvalues with real parts less than or equal to zero as opposed to strictly less than), and using center manifold theory to prove stability along the zero eigenvalue manifolds. This modification would also require new guidelines for allowable choices of  $h$ , namely, penalizing only powers of the state affectable through  $A$  or  $B$ . This conjecture is based on the analytical solu-*

tion results of Chapter 3 and the scalar system results of Chapter 5, and has been verified in a small number of simulations.

### 12.5 Addition of Off-axis Rotor for Stabilizability/Detectability

We saw in the previous section that the SDRE regulator cannot in general achieve asymptotic stability of the desired equilibrium point for the baseline gyrostat. We therefore alter the base configuration by addition of a second, off-axis rotor with control law already specified to achieve energy dissipation or damping in the closed loop system. The effect of this damping is achieved by introduction of an additional, coupled state equation for the off-axis rotor angular momentum about its spin axis, which we shall call  $\mu_2$ . Hall has developed and analyzed the resulting equations of motion for the two rotor case [24], and we give them here for the specific choice of the off-axis rotor spin axis being in purely the  $e_2$  direction. For the reduced system with  $x_1$  a function of  $q$  (recall (12.12)) we get

$$\begin{aligned}\dot{q}_1 &= (\nu - i_2 x_1) q_2 + q_2 - \alpha_2 x_1 \mu_2 \\ \dot{q}_2 &= (i_3 x_1 - \nu) q_1 - q_1 \\ \dot{\mu}_2 &= e_2 q_2 - d_2 \mu_2 \\ \dot{\nu} &= u\end{aligned}\tag{12.26}$$

where

$$\begin{aligned}\alpha_2 &= \frac{1}{I_2} \\ e_2 &= \alpha_2 I_{s_2} \\ d_2 &= 1 + e_2\end{aligned}\tag{12.27}$$

and  $I_{s_2}$  is the off-axis rotor principal moment of inertia with respect to the  $e_2$  axis. For (12.26) we now find the Jacobian at zero to be

$$J(0) = \begin{bmatrix} 0 & 1 - i_2 & -\alpha_2 & 0 \\ i_3 - 1 & 0 & 0 & 0 \\ 0 & e_2 & -d_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (12.28)$$

so that the linearized controllability matrix is again given by

$$M_{cf}(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (12.29)$$

but now stabilizability of the linearization requires that the eigenvalues of the upper left  $3 \times 3$  subblock of (12.28) have negative real parts. If one realizes the simple equalities  $1 - i_2 = \alpha_2$  and  $i_3 - 1 = -1/I_3 \equiv -\alpha_3$ , then the characteristic equation of the said subblock is easily computed to be

$$\lambda^3 + d_2\lambda^2 + \alpha_2\alpha_3\lambda + \alpha_2\alpha_3 = 0 \quad (12.30)$$

Now, using Routh's criterion [17], we compute the Routhian array

$$\begin{array}{cc} 1 & \alpha_2\alpha_3 \\ d_2 & \alpha_2\alpha_3 \\ (d_2\alpha_2\alpha_3 - \alpha_2\alpha_3)/d_2 & \\ \alpha_2\alpha_3 & \end{array} \quad (12.31)$$

and left half plane eigenvalues of (12.28) are guaranteed if all entries in the first column of (12.31) have the same sign. Since the third element in the first column of (12.31) simplifies to  $e_2\alpha_2\alpha_3/d_2$  and since  $\alpha_2$ ,  $\alpha_3$ ,  $d_2$ , and  $e_2$  are all positive, we satisfy Routh's criterion so that the linearization of (12.26) is indeed stabilizable. We also see from this analysis that detectability will be guaranteed, provided  $\nu$  has an independent, globally positive definite penalty via the SDC parametrization  $h(x) = H(x)x$ .

The above discussion shows that adding the off-axis damping rotor guarantees stabilizability and detectability in some small neighborhood of the origin, for *any* parametrization  $A$ , and for a suitable parametrization  $H$  which has a positive definite term independently penalizing  $\nu$ . We are interested, however, in performing transverse spinup maneuvers which start far away from the origin, so that something better than stabilizability/detectability of the linearization is desired. We thus propose a factorization  $A$  which guarantees stabilizability/detectability everywhere in the positive momentum sphere ( $x_1 > 0$ ). We first make some general comments regarding factorizations. Although it is always possible to introduce fictitious terms of second and higher order into the dynamics which theoretically cancel, in numerical implementations it is unlikely that exact cancellations will result, and, even if they do cancel, adding such terms may not be a good idea. To illustrate, for the two state system

$$\begin{aligned}\dot{x}_1 &= a_1(x) \\ \dot{x}_2 &= a_2(x)\end{aligned}\tag{12.32}$$

suppose that  $a_1$  and  $a_2$  are purely linear functions of  $x$ . Then a natural choice for  $A(x)$  is the actual matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

that results from evaluating the Jacobian of  $a$  at  $x = 0$ . However, strictly speaking

$$A(x) = \begin{bmatrix} a_{11} + c_1 x_2 & a_{12} - c_1 x_1 \\ a_{21} + c_2 x_2 & a_{22} - c_2 x_1 \end{bmatrix}\tag{12.33}$$

is a valid parametrization for any  $c_1$  and  $c_2$ , since the  $c_i x_1 x_2$  terms cancel in  $A(x)x$ . The addition of these fictitious factors, however, changes the pointwise appearance of the dynamics from the true dynamics given by the matrix  $A$ . Thus, although such techniques have been used for  $a = 0$  to derive stable control laws using the SDRE technique [14], they will not be used here.

Given this consideration, when we consider (12.26) and realize that  $x_1$  is no longer considered a state nor easily factored into linear functions of the states  $q_1$  and  $q_2$ , then we see that the only

terms we need to decide how to factor are the  $\nu q_2$ ,  $\nu q_1$  terms in the  $q_1$ ,  $q_2$  dynamics, respectively. Since  $\nu$  is completely controllable, heuristically we want the factorization to show strong pointwise linear controllability of  $q_1$  and  $q_2$  through  $\nu$ . We thus propose the factorization

$$A(x) = \begin{bmatrix} 0 & 1 - i_2 x_1 & -\alpha_2 x_1 & q_2 \\ i_3 x_1 - 1 & 0 & 0 & -q_1 \\ 0 & e_2 & -d_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (12.34)$$

This factorization has the factored controllability matrix function

$$M_{cf}(x) = \begin{bmatrix} 0 & q_2 & -\xi_2 q_1 & \xi_1 \xi_2 q_2 + \alpha_2 e_2 x_1 q_1 \\ 0 & -q_1 & \xi_1 q_2 & -\xi_1 \xi_2 q_1 \\ 0 & 0 & -e_2 q_1 & e_2 (\xi_1 q_2 + d_2 q_1) \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (12.35)$$

where

$$\begin{aligned} \xi_1 &= i_3 x_1 - 1 \\ \xi_2 &= 1 - i_2 x_1 \end{aligned} \quad (12.36)$$

The matrix function (12.35) has determinant

$$\det[M_{cf}] = -e_2 [\xi_1^2 q_2^3 + d_2 \xi_1 q_2^2 q_1 - \xi_1 \xi_2 q_2 q_1^2 + (e_2 \alpha_2 x_1 - d_2 \xi_2) q_1^3] \quad (12.37)$$

which obviously loses rank for nontrivial values of  $q_1$  and  $q_2$ , so that (12.34) does not yield a globally controllable parameterization. However, (12.34) does yield guaranteed stabilizability for  $x_1 > 0$ , as we now show by Routh's criterion and some simple analysis. For stabilizability of  $\{A, B\}$ , recall that we must have [76] that, for all  $\lambda$  and  $y$  such that  $y^T A = y^T \lambda$  and  $\operatorname{Re} \lambda \geq 0$ ,  $y^T B \neq 0$ . Thus, we consider the eigenstructure of (12.34). Setting the determinant of  $A(x) - \lambda I$  equal to zero we find the characteristic equation

$$\lambda(\lambda^3 + d_2 \lambda^2 - \xi_1 \xi_2 \lambda + \xi_1 \xi_3) = 0 \quad (12.38)$$

where  $\xi_1, \xi_2$  are as in (12.36) and

$$\xi_3 = \alpha_2 e_2 x_1 - \xi_2 d_2 \quad (12.39)$$

We thus have a single zero eigenvalue, plus three more eigenvalues determined by the roots of the term in parentheses in (12.38). For the zero eigenvalue, we easily find the corresponding left eigenvector to be  $y^T = [0 \ 0 \ 0 \ 1]$ . Since  $y^T B = 1 \neq 0$ , this zero eigenvalue is stabilizable. We now show that the remaining three eigenvalues have negative real parts under some slight additional assumptions, so that stabilizability is guaranteed. Recall that a necessary condition for only left half plane roots of a polynomial is that all the coefficients have the same sign [17]. We thus must have

$$d_2 > 0 \quad (12.40)$$

$$-\xi_1 \xi_2 > 0 \quad (12.41)$$

$$\xi_1 \xi_3 > 0 \quad (12.42)$$

We have (12.40) satisfied trivially, and since  $x_1 \leq 1$  and  $i_2, i_3 < 1$ , we also have from (12.36) that  $\xi_1 < 0$  and  $\xi_2 > 0$ . Thus, (12.41) is satisfied, while for satisfaction of (12.42) we require  $\xi_3 < 0$ . We must now break up the analysis into prolate and oblate cases.

For the prolate case recall we have  $i_2 > i_3 > 0$ . Thus, from (12.36) we see in this case that, as long as  $x_1 > 0$ , we have

$$\xi_3 < \alpha_2 e_2 + (i_2 - 1)d_2 \quad (12.43)$$

Using the definition of  $i_2$  (12.1) and  $\alpha_2$  (12.27), (12.43) becomes

$$\xi_3 < \alpha_2 e_2 - \alpha_2 I_p d_2 = \alpha_2 (e_2 - I_p d_2) \quad (12.44)$$

Now, using the definitions of  $d_2$  and  $e_2$  (12.27), (12.44) can be written

$$\xi_3 < \alpha_2 (\alpha_2 I_{s_2} (1 - I_p) - I_p) \quad (12.45)$$

Thus,  $\xi_3 < 0$  and we satisfy the necessary condition for left half plane eigenvalues regardless of the values of  $\alpha_2$  and  $I_{s_2}$  if  $I_p \geq 1$ . In the examples done in this chapter, we have nondimensionalized the problem so that  $I_p = 1$ , and thus we satisfy the necessary condition. Note that even if  $0 < I_p < 1$ ,

then the off-axis rotor moment of inertia  $I_{s_2}$  is typically quite small compared to  $I_p$ , so that the necessary condition should be satisfied.

Returning now to the oblate case, recall we have  $i_3 < i_2 < 0$ . Thus, from (12.36) we see in this case that, as long as  $x_1 > 0$ , we have

$$\xi_3 < \alpha_2 e_2 - d_2 \quad (12.46)$$

Now, using the definitions of  $d_2$  and  $e_2$  (12.27), (12.46) can be written

$$\xi_3 < I_{s_2} \alpha_2 (\alpha_2 - 1) - 1 \quad (12.47)$$

Thus,  $\xi_3 < 0$  and we satisfy the necessary condition for left half plane eigenvalues in the oblate case regardless of the value of  $I_{s_2}$  if  $\alpha_2 < 1$ . Even if this is not the case, the smallness of  $I_{s_2}$  is usually sufficient to guarantee  $\xi_3 < 0$ . For example, in this chapter the simulations for the oblate case assume  $I_{s_2} = 0.1$  and  $\alpha_2 = 1.3$ , so that (12.47) gives  $\xi_3 < -0.961$ .

Thus, we have necessary conditions for stability of the three nonzero eigenvalues of (12.34), and we have shown that we satisfy these conditions for the examples simulated in this dissertation. We still need to demonstrate satisfaction of a sufficient condition for stability of these eigenvalues, however. We now do so by computing the Routhian array for the parenthetical term in (12.38) to find

$$\begin{array}{cc} 1 & -\xi_1 \xi_2 \\ d_2 & \xi_1 \xi_3 \\ (-d_2 \xi_1 \xi_2 - \xi_1 \xi_3)/d_2 & \\ \xi_1 \xi_3 & \end{array} \quad (12.48)$$

and observing that for sufficiency we thus need to satisfy the necessary conditions (12.40) and (12.42), along with the additional condition

$$-\xi_1(d_2 \xi_2 + \xi_3) > 0 \quad (12.49)$$

Since  $\xi_1 < 0$ , using the definition of  $\xi_3$  we find that (12.49) becomes

$$d_2 \xi_2 + \alpha_2 x_1 e_2 - \xi_2 d_2 = \alpha_2 x_1 e_2 > 0 \quad (12.50)$$

which holds as long as  $x_1 > 0$ , which has been assumed throughout. Thus, we have shown that (12.34) gives a stabilizable parametrization for both the oblate and prolate cases, as long as trajectories remain in the positive ( $x_1$ ) momentum sphere.

**Remark:** *The existence of this stabilizable parametrization enabling use of the SDRE technique throughout the positive hemisphere is very interesting, particularly for the prolate case, since a manifold of open loop equilibrium points with unstabilizable local linearizations is known to exist [24] in the positive hemisphere for the prolate case. The implication here is that, if a trajectory were to pass through this manifold, then a control algorithm based on Riccati equations and local linearizations such as the LMI based techniques of [5] and [74] would fail to yield computable controllers at such a point, whereas the SDRE methods, which use the exact parametrization (12.34) of the dynamics instead of a local linearization, still work satisfactorily at such points.*

Note that the parametrization (12.34) corresponds to the choices of  $c_1 = c_2 = 0$  for the more general possible factorization

$$A(x) = \begin{bmatrix} 0 & 1 - i_2 x_1 + c_1 \nu & -\alpha_2 x_1 & (1 - c_1) q_2 \\ i_3 x_1 - 1 - c_2 \nu & 0 & 0 & -(1 - c_2) q_1 \\ 0 & e_2 & -d_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (12.51)$$

and we observe that, if we pick  $c_1 = c_2 = 1$ , then the  $q$  dynamics appear to be pointwise unaffected from the controllable variable  $\nu$ , so that the factored controllability matrix for such a system is equal to (12.29) for all  $x$ . Thus, the uncontrollable space for (12.51) from a factorization standpoint is all of the  $(q_1, q_2, \mu_2)$  space, as opposed to the much smaller subset of this space determined by setting (12.37) equal to zero, and determining the left nullspace of (12.35) at the resulting points. Stabilizability of the factorization (12.51) thus requires that all three nonzero eigenvalues of (12.51) be stable, which was sufficient but far from necessary for stabilizability of the chosen parametrization

(12.34). Computing the characteristic equation of (12.51) we find

$$\lambda[\lambda^3 + d_2\lambda^2 + (\nu - \xi_1)(\xi_2 + \nu)\lambda + (\xi_1 - \nu)(e_2\alpha_2x_1 - d_2(\xi_2 + \nu))] = 0 \quad (12.52)$$

and we see that (12.52) differs from (12.38) only in the first and zeroth-order coefficients of  $\lambda$ . However, this difference is significant since, if we consider the prolate case with  $x_1 \approx 1$ , there exist values of  $\nu$  which render the first-order coefficient negative. For example, let  $i_2 = 0.5$  and  $i_3 = 0.3$ . Then the first-order coefficient becomes  $(\nu+0.7)(\nu+0.5)$ , so that for all values of  $\nu \in (-0.5, -0.7)$ , the coefficient is negative, and thus the factorization cannot be stabilizable. As an additional comment we mention that this issue is not just of theoretical importance for this problem. Simulated attempts at SDRE nonlinear regulation of the gyrostat using (12.51) became numerically unstable when points in the trajectory were encountered where the factorization was not stabilizable. We mention this case in particular because it corresponds to a natural SDC parametrization that arises from considering the gyrostat as a Hamiltonian system as in [23]. In this framework, the nonreduced equations of motion for the gyrostat can be written

$$\begin{aligned} \dot{x} &= -\nabla \mathcal{H}^\times x \\ \dot{\mu} &= u \end{aligned} \quad (12.53)$$

where  $\nabla$  represents the gradient operator and  $\mathcal{H}$  is an appropriate Hamiltonian. Thus, the point to be made from all the above analysis is that selection of an appropriate SDC factorization for a nonlinear system can be nontrivial, and pointwise controllability issues should play a strong role, whereas other issues such as elegance of derivation of an SDC form may be inconsequential.

Finally, we need to consider detectability issues. Recall that for detectability we must have that  $Ay = \lambda y$  and  $\operatorname{Re} \lambda \geq 0$  implies  $Hy \neq 0$ . Now, since we only have one eigenvalue of (12.34) that has  $\operatorname{Re} \lambda \geq 0$  ( $\lambda = 0$ ), then the requirements for detectability are the same as the necessary condition for global asymptotic stability given in Chapter 9, namely that if  $Ay = 0$ , then we must have  $Hy \neq 0$  for  $y \neq 0$ . A sufficient condition for global detectability is clearly then to pick a globally nonsingular

$H$ , as pointed out in Chapter 9. We therefore propose a choice of  $H$  to be

$$H = \text{diag}[k_1, k_2, k_3, k_4] \quad (12.54)$$

with  $k_i > 0$  for all  $i = 1, 2, 3, 4$ . Also, as long as  $k_4 \neq 0$ , then  $Hy = 0$  implies  $\nu = 0$ . For  $\nu = 0$ , then  $\dot{q}_2 = a_2 = 0$  implies  $q_1 = 0$ . Finally, setting  $a_1 = a_3 = 0$  gives

$$\begin{aligned} q_2 &= \frac{\alpha_2 x_1}{\xi_2} \mu_2 \\ q_2 &= \frac{d_2}{e_2} \mu_2 \end{aligned} \quad (12.55)$$

which, unless  $\alpha_2 x_1 / \xi_1 = d_2 / e_2$ , are lines that intersect only at  $q_2 = \mu_2 = 0$ . Thus, for nonequality of the two slopes in (12.55),  $Hy \neq 0$  if  $\nu \neq 0$ , and if  $\nu = 0$  then  $Ay = 0$  only for  $y = 0$ . Substituting the parameter values used later in the simulations, for the oblate case we get equality of the slopes for  $x_1 = -6.65$ , and we get  $x_1 = 1.79$  for the prolate case. Since neither of these  $x_1$  values is possible, taking  $k_1 = k_2 = k_3 = 0$  and  $k_4 \neq 0$  in (12.54) guarantees global detectability and also satisfaction of the necessary condition for global asymptotic stability of the closed loop system. We investigate the effect of using these two choices for  $H$  in the next section.

## 12.6 SDRE Nonlinear Regulator Simulation Results

In this section we give typical simulation results for the three spinup maneuvers described in Section 12.1 using the sampled data SDRE nonlinear regulator with  $A(x)$  and  $H(x)$  selected according to (12.34) and (12.54), respectively. For comparison purposes we also include the same maneuvers performed under the commonly made [23] small, constant torque assumption. It is shown that the SDRE nonlinear regulator is more effective in driving the cone angle to zero than the constant torque maneuver, and also offers significant design flexibility while yielding closed loop stability. All simulations were performed in Matlab/Simulink using Runge-Kutta, fourth-order numerical integration.

Prior to giving the results for the gyrostat with off-axis rotor, we first show some typical simulation results for the oblate spinup of the four-state gyrostat (three state design model (12.23))

without damping. Recall that this model does not have a stabilizable linearization at the origin, but instead has  $q_1(x_3)$  and  $q_2(x_2)$  states with purely imaginary eigenvalues, and completely controllable  $\nu$  mode with zero eigenvalue. Results are given for two choices of  $H$ , as discussed at the end of the previous section. In both simulations the sampling rate was 10 Hz, arbitrarily chosen to be an implementable number which yielded smooth-looking trajectories on the time scale of the problem. The integration step size was an order of magnitude smaller, at 0.01 seconds. The initial condition was  $(q_1, q_2, \nu) = (\sin(10^\circ), 0, 0.2 \cos(10^\circ) - 1)$ , corresponding to the all-spun condition with an initial cone angle of ten degrees and  $i_1 = 0.2$ . In all simulations in this section, we fix the control penalty weight  $R$  at the value one, and for all the oblate spinup maneuvers in this section, we use the additional inertia parameters  $i_2 = -0.3, i_3 = -0.5$ . The results for

$$H = H_1 = \text{diag}(0, 0, 0.1)$$

are shown in Figure 12.2. Note that the penalty value of 0.1 on  $\nu$  was selected to keep the control input in a neighborhood of 0.1 or less, which is known [23] to be a reasonable control magnitude for oblate spinup maneuvers. As expected for this choice of  $H$ ,  $\nu$  is driven to zero, but  $q_1(x_3)$  and  $q_2(x_2)$  enter into a limit cycle due to their pure imaginary axis eigenvalue pair. Note that in order to actually solve the SDRE for this problem, we need solve only the scalar SDRE

$$-p^2 + 0.01 = 0$$

which yields the linear control

$$u = -0.1\nu$$

This problem cannot be solved using a three-state parametrization for  $A(x)$  because of the lack of detectability.

The results for

$$H = H_2 = \text{diag}(1, 1, 0.1)$$

are shown in Figure 12.3. Note that global observability is guaranteed by the global nonsingularity of  $H$ . The factorization used for this simulation is the same as (12.34), except the  $\mu_2$  state has been

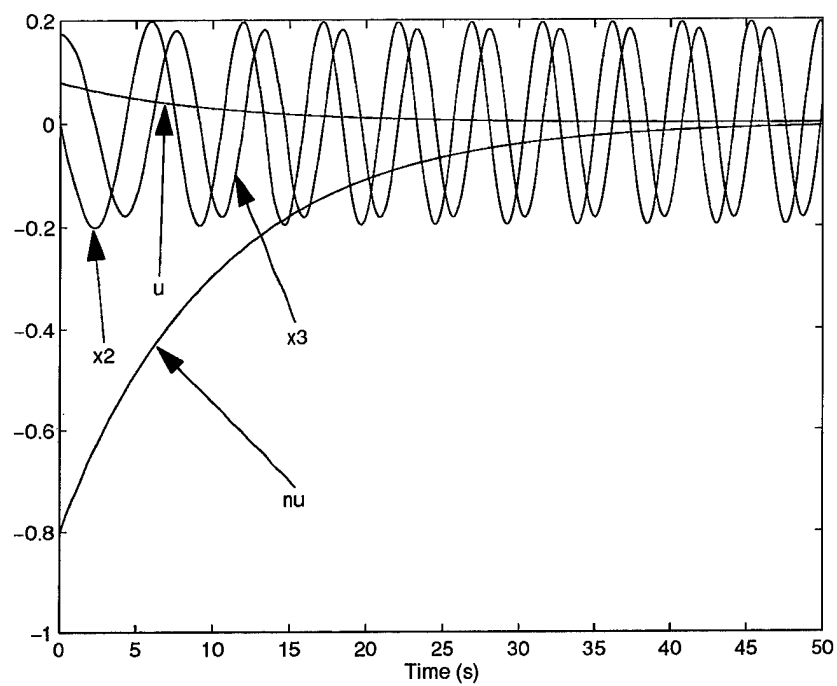


Figure 12.2: SDRE State and Control Histories for 4-State Gyrostat ( $H = H_1$ )

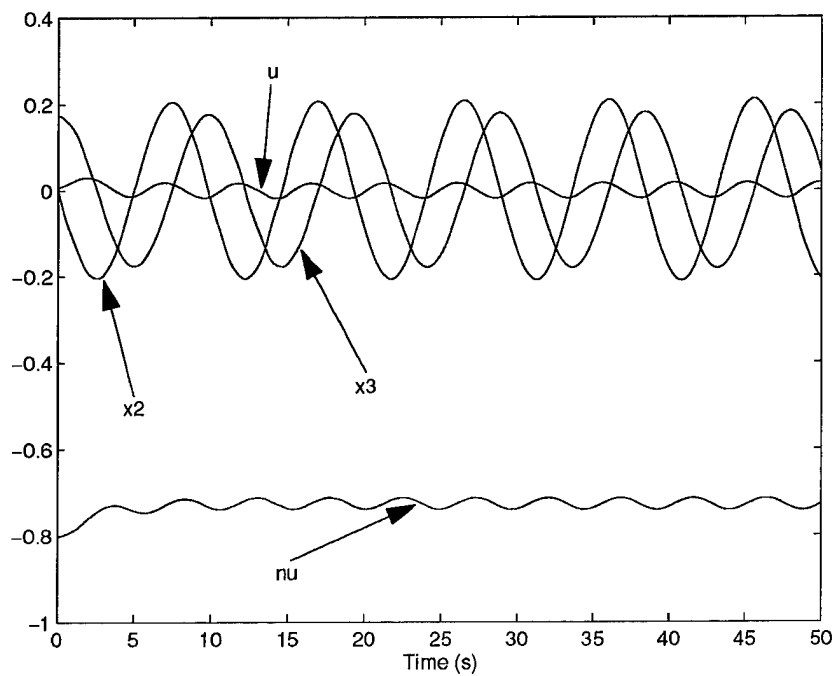


Figure 12.3: SDRE State and Control Histories for 4-State Gyrostat ( $H = H_2$ )

removed. Thus,

$$A(x) = \begin{bmatrix} 0 & 1 - i_2 x_1 & q_2 \\ i_3 x_1 - 1 & 0 & -q_1 \\ 0 & 0 & 0 \end{bmatrix} \quad (12.56)$$

for this example. It is easily verified that (12.56) gives a controllable parametrization as long as  $q_1$  and  $q_2$  are not both zero. This lack of controllability at the origin, however, prevents asymptotic stability, as seen in the figure. Interestingly,  $\nu$  is not regulated much toward zero in this case, for which  $q_1$  and  $q_2$  are being penalized ten times as much as  $\nu$ . For fixed unit penalties on  $q_1$  and  $q_2$ , as the penalty on  $\nu$  was increased, additional simulations showed final values of  $\nu$  increasingly closer to zero.

Finally, we compare these results to the constant torque case with  $u = 0.02$ . The time histories for this case are shown in Figure 12.4. Note that in all cases the lack of a stabilizable linearization at the origin prevents obtaining asymptotic closed loop stability. As a final basis of comparison, we show the  $x_1$  time histories for the three approaches in Figure 12.5. In this and succeeding figures,  $x1c$  is the constant torque history,  $x1s1$  is the SDRE history for  $H = H_1$ , and  $x1s2$  is the SDRE history for  $H = H_2$ . As expected, none of the  $x_1$  trajectories asymptotically approach the value one as desired. Also, the SDRE case for nonsingular  $H$  has the worst performance in terms of driving  $x_1$  to the desired nonzero setpoint, although it is not much worse than the others in an absolute sense.

We now present simulation results for the five-state gyrostat model, which includes the off-axis rotor to provide damping and thus a stable linearization at the origin. We give typical results for oblate, prolate, and transverse spinup maneuvers, in that order. Recall that for this system we actually use a four-state design model in the SDRE methods, so that we use the  $A(x)$  parametrization given by (12.34) for all maneuvers. We again investigate the results of different choices for  $H$ , although we restrict  $H$  to the diagonal, constant form given by (12.54). Starting with the oblate case, we again show results for simulations run at 10 Hz sampling rate, with inertia parameters as above. From now on we use  $x_3$  and  $x_2$  when referring to  $q_1$  and  $q_2$ , respectively, so that for these simulations we start with initial conditions of  $(x_3, x_2, \mu_2, \nu) = (\sin(15^\circ), 0, 0, 0.2 \cos(15^\circ) - 1)$ ,

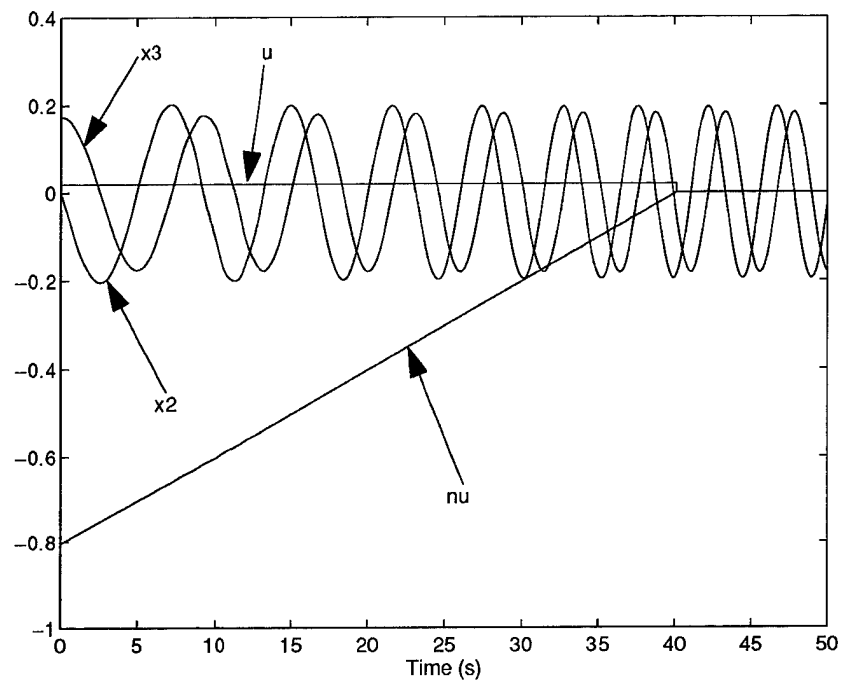


Figure 12.4: State and Control Histories for 4-State Gyrostat ( $u = 0.02$ )

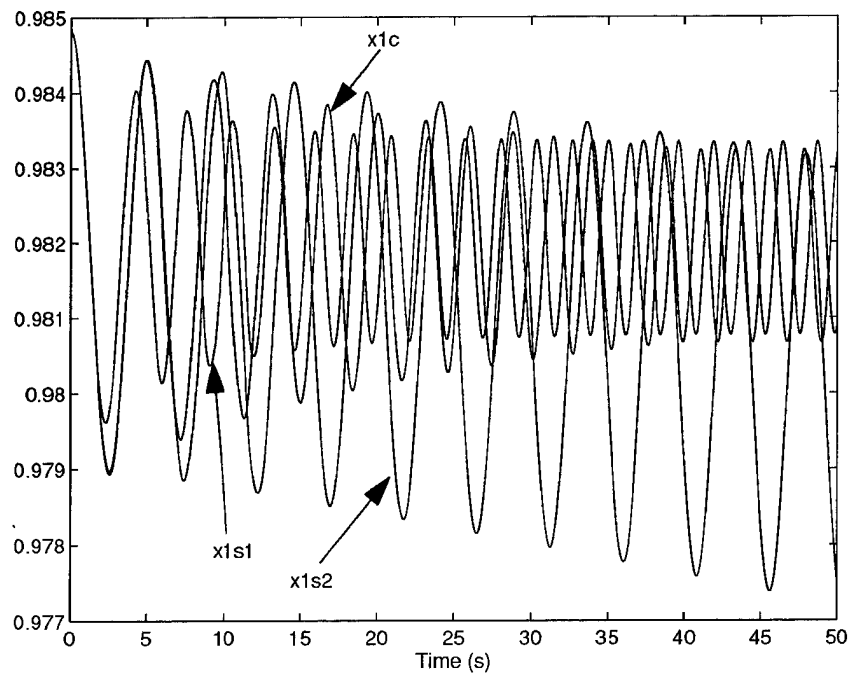


Figure 12.5:  $x_1$  State Histories for 4-State Gyrostat

corresponding to the all-spun condition with an initial cone angle of  $15^\circ$ . In computing the  $A(x)$  factorization, as mentioned before, we have selected  $I_p = 1$  so that for the oblate case we obtain  $\alpha_2 = 1.3$  and  $\alpha_3 = 1.5$ . We have also selected  $I_{s_2} = 0.1$  so that  $e_2 = 0.13$  and  $d_2 = -1.13$  in (12.34). We show state history results for

$$H = H_1 = \text{diag}(0, 0, 0, 0.1) \quad (12.57)$$

in Figure 12.6, and for

$$H = H_2 = \text{diag}(1, 1, 1, 0.1) \quad (12.58)$$

in Figure 12.7. Note Figure 12.7 has a longer time scale to better illustrate convergence of  $\nu$ .

For purposes of comparison, we also show state histories for a constant applied torque of  $u = 0.01$  in Figure 12.8. Note that even though the  $\nu$  histories are quite different for the various simulations (in particular  $\nu$  is driven to zero much faster for  $H = H_1$ ), there is little difference between the  $x_2$ ,  $x_3$ , and  $\mu_2$  histories. A comparison of the three controls and  $x_1$  histories are given in Figures 12.9 and 12.10. In the figures, as before, the  $c$  suffix denotes the constant torque case, whereas  $s1$  corresponds to SDRE with  $H = H_1$ , and  $s2$  corresponds to SDRE with  $H = H_2$ . Figure 12.9 is interesting in that the three controls show significantly different natures, leading to the different  $\nu$  histories observed, but the other state histories are quite similar. Looking at Figure 12.10 we see that SDRE with  $H = H_2$  does a slightly better job than the other methods of driving  $x_1$  to one, but there is not much difference between any of the methods. Conceptually, this makes sense since the ( $H = H_2$ ) SDRE case is actively trying to drive  $x_2$  and  $x_3$  to zero, whereas the ( $H = H_1$ ) SDRE and constant torque cases ignore the other states and simply work on regulating  $\nu$ .

Before studying the prolate case, we remark that the results observed here agree with the analysis of Hall [24] for the oblate case, in that we expect the small, constant torque to be effective in performing the oblate spinup maneuver, since no open loop separatrices are crossed along the trajectory. In the prolate case, which we now address, we do however expect the small constant torque solutions to show an increased cone angle, and thus the SDRE methods may offer some advantages.

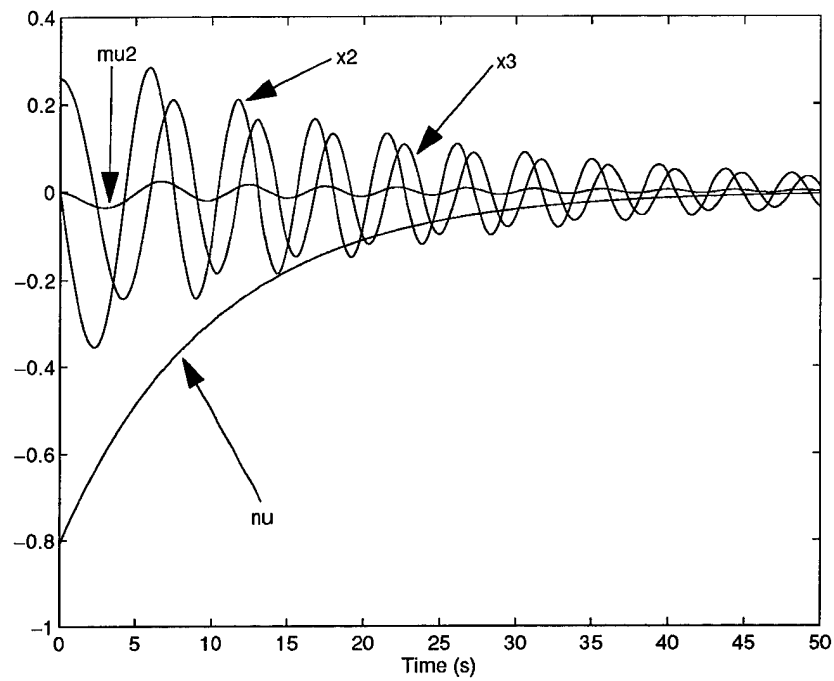


Figure 12.6: Oblate SDRE State Histories for 5-State Gyrostat ( $H = H_1$ )

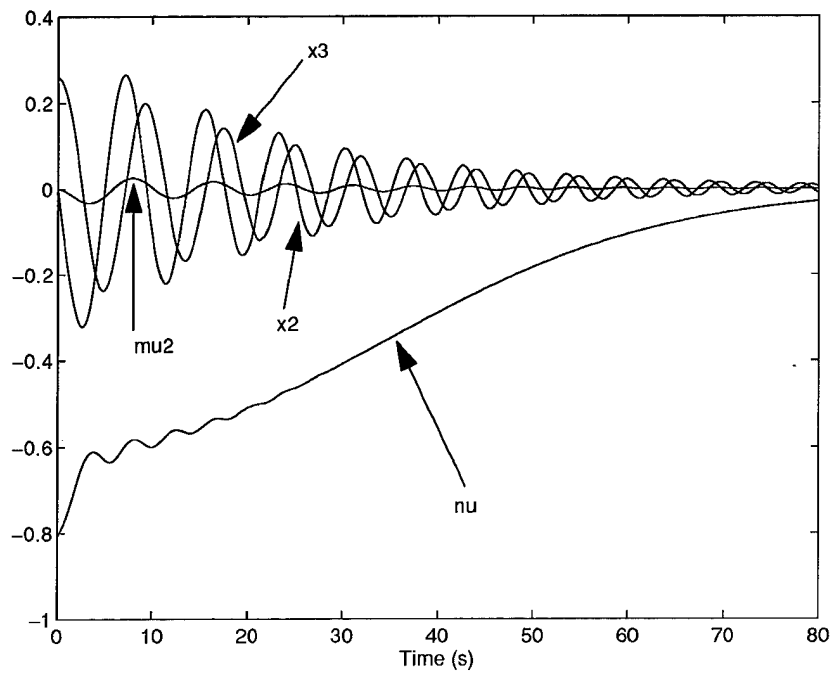


Figure 12.7: Oblate SDRE State Histories for 5-State Gyrostat ( $H = H_2$ )

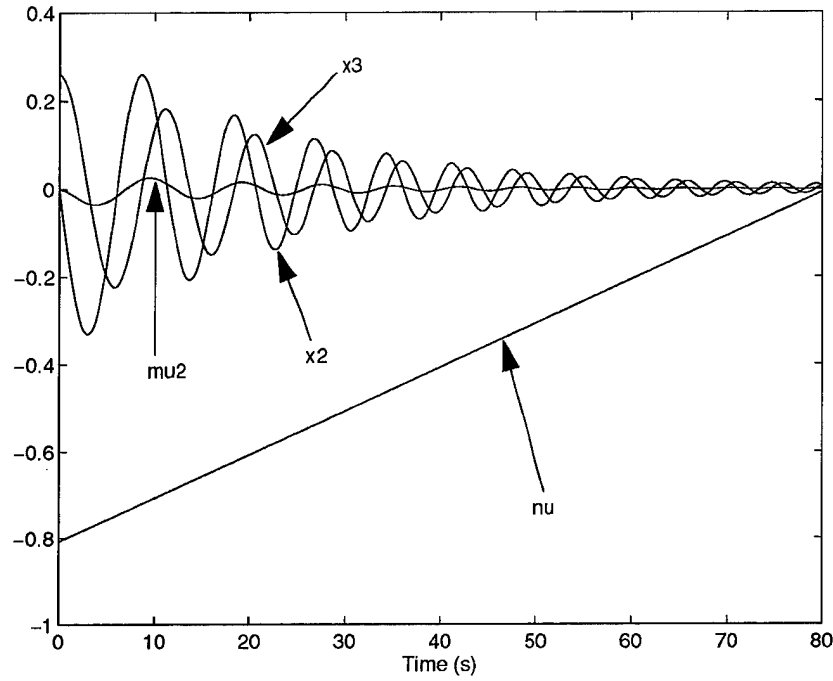


Figure 12.8: Oblate State Histories for 5-State Gyrostat ( $u = 0.01$ )

For the prolate simulations, we use the same equations and factorizations as for the oblate case, but we now change our inertia parameters to  $i_2 = 0.5$  and  $i_3 = 0.3$ , with  $i_1 = 0.2$  as before. This leads to  $\alpha_2 = 0.5$  and  $\alpha_3 = 0.7$ . By changing  $I_{s_2}$  to 0.26, we maintain the factorization parameters  $e_2$  and  $d_2$  the same as in the oblate case. To facilitate comparison between cases, we start the prolate simulations from the same initial condition as was used in the oblate simulations, and we again sample at 10 Hz. This time we start by examining the small constant torque case. In Figure 12.11 we give state time histories for  $u = 0.01$  until  $\nu = 0$ , and then  $u = 0$  thereafter. Two things are immediately apparent from the figure. Note first of all that changing to the prolate case has significantly increased the settling time of the  $(x_2, x_3, \mu_2)$  states over the oblate case. This is because in the oblate case, the eigenvalues corresponding to these states in the linearization about the origin are  $\lambda = -0.044 \pm 1.46j$ ,  $-1.042$ , respectively, whereas in the prolate case they become  $\lambda = -0.015 \pm 0.56j$ ,  $-1.1$ . Thus, the oscillatory modes in the prolate case have a much slower decay rate. The second thing to notice is the large amplitude oscillations in the  $x_2$  and  $x_3$  states. Since

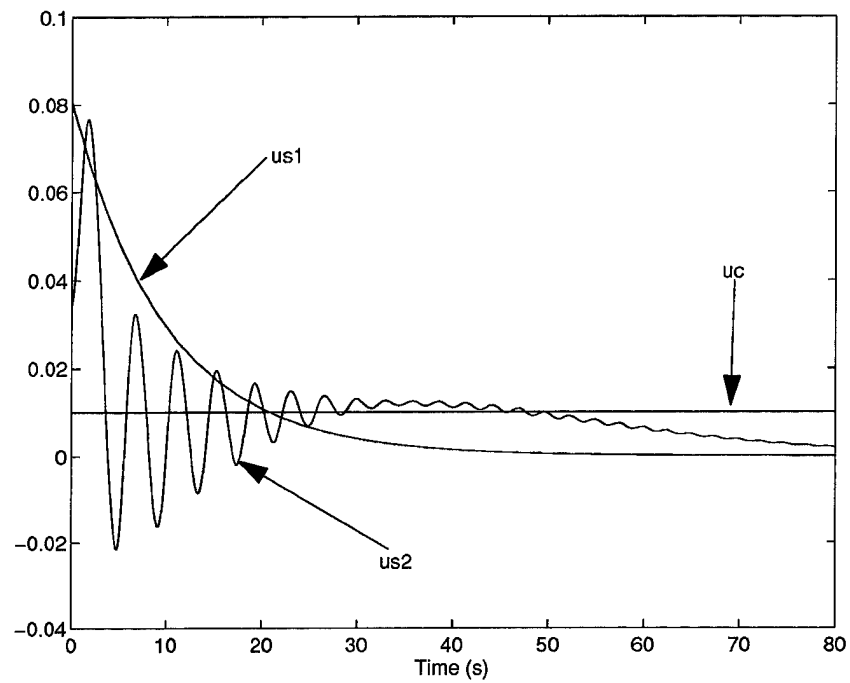


Figure 12.9: Oblate Control Histories for 5-State Gyrostat

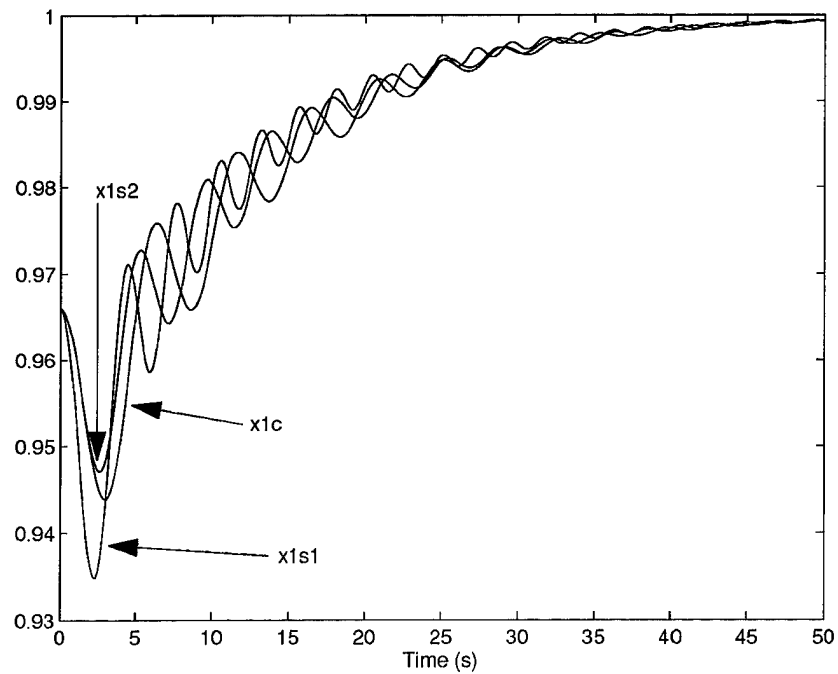


Figure 12.10: Oblate  $x_1$  Histories for 5-State Gyrostat

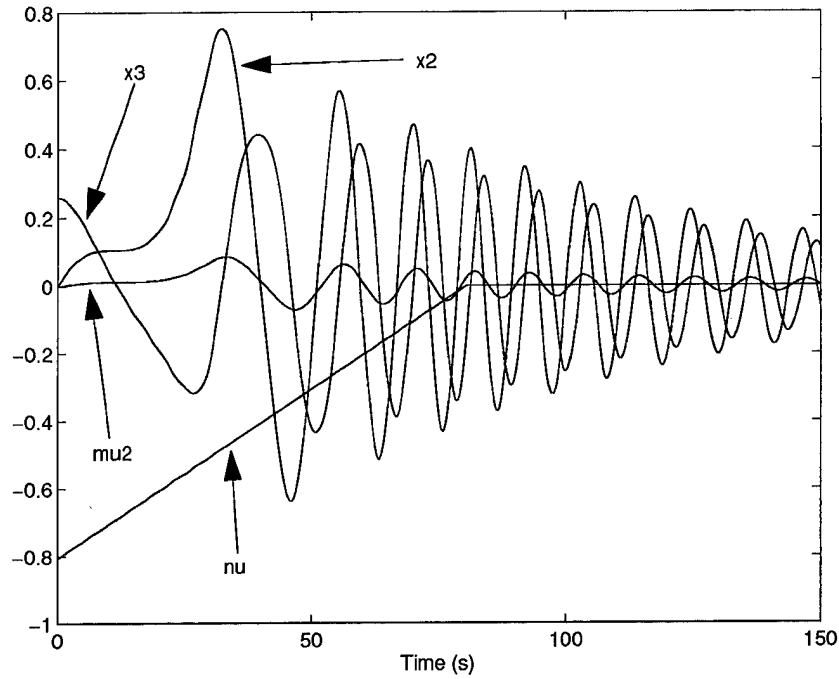


Figure 12.11: Prolate State Histories for 5-State Gyrostat ( $u = 0.01$ )

our control strategy is concerned only with driving  $\nu$  to zero, these states range far from our actual desired values of zero.

In Figures 12.12 and 12.13, we show state histories for the SDRE regulators with  $H = H_1$  and  $H = H_2$ , respectively. We again see the longer time scale for the prolate case, but note that the SDRE regulators do not allow nearly as large oscillations of the  $x_2$  and  $x_3$  states, as reflected in the plots' smaller vertical scale. This should correspond to better performance of these controllers in driving  $x_1$  to one. In Figures 12.14 and 12.15, we compare control and  $x_1$  time histories for the three methods. Indeed, we see that the SDRE methods perform better than the constant torque controller in driving  $x_1$  to one, with the  $H = H_2$  controller doing the best job. The price to be paid is seen in Figures 12.13, in which regulation of  $\nu$  is seen to take much longer for the  $H = H_2$  controller than for the other control strategies, and in Figure 12.14, in which we see the SDRE  $H = H_2$  controller applies a comparatively large magnitude initial value of the control.

We conclude this section with results for the transverse spinup maneuver, assuming the same inertia values used for the prolate spacecraft just examined. A sampling rate of 10 Hz was again

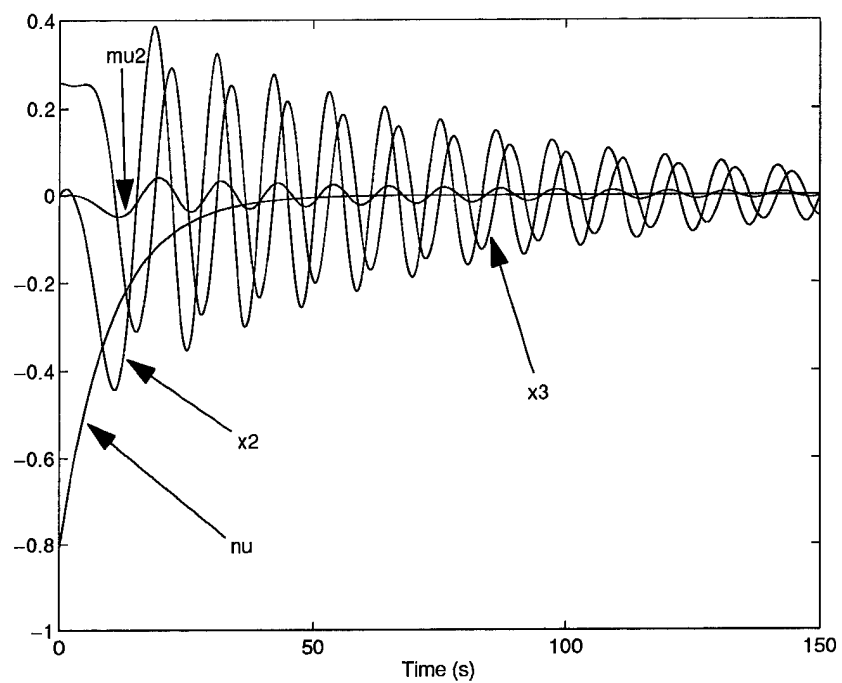


Figure 12.12: Prolate SDRE State Histories for 5-State Gyrostat ( $H = H_1$ )

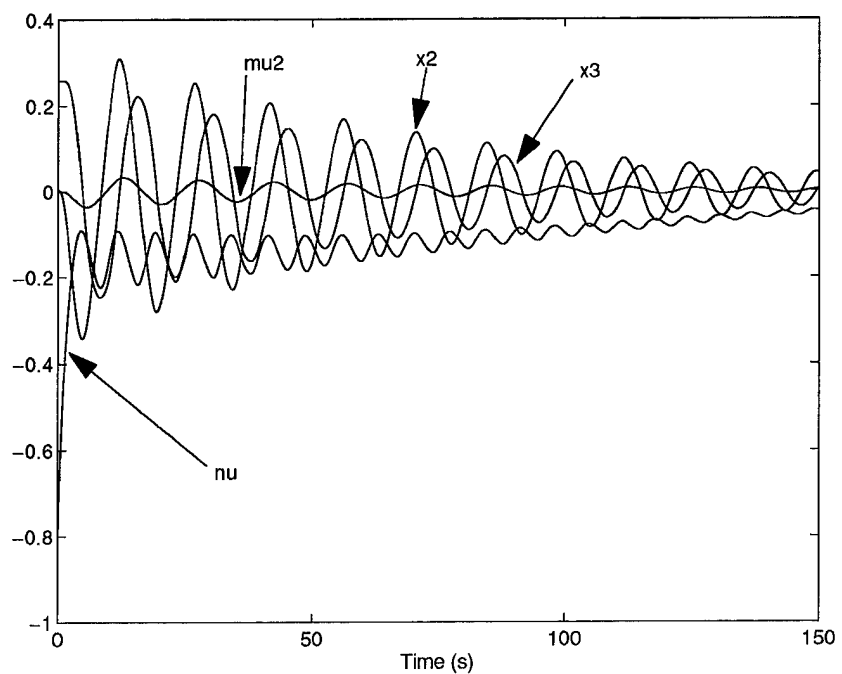


Figure 12.13: Prolate SDRE State Histories for 5-State Gyrostat ( $H = H_2$ )

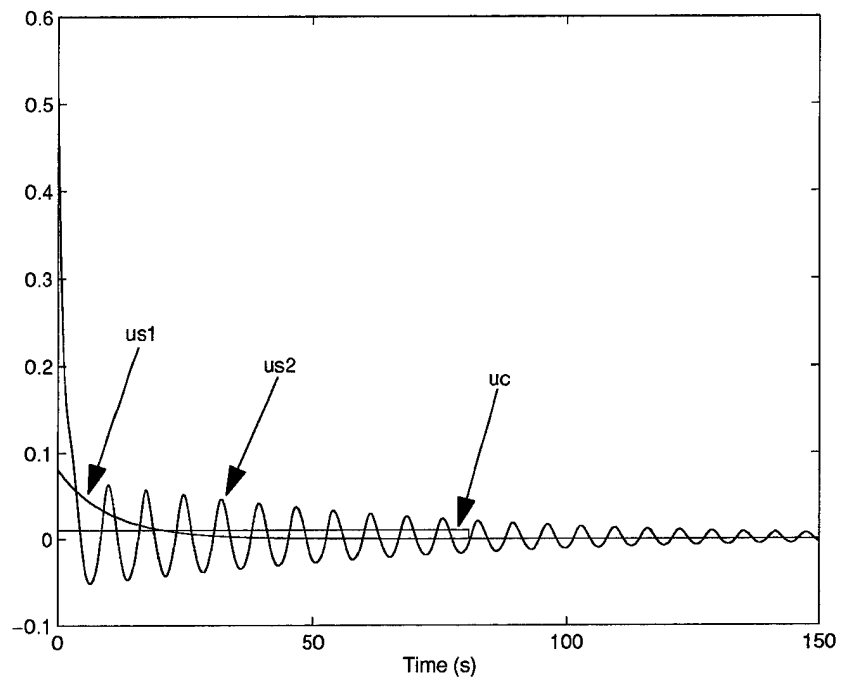


Figure 12.14: Prolate Control Histories for 5-State Gyrostat

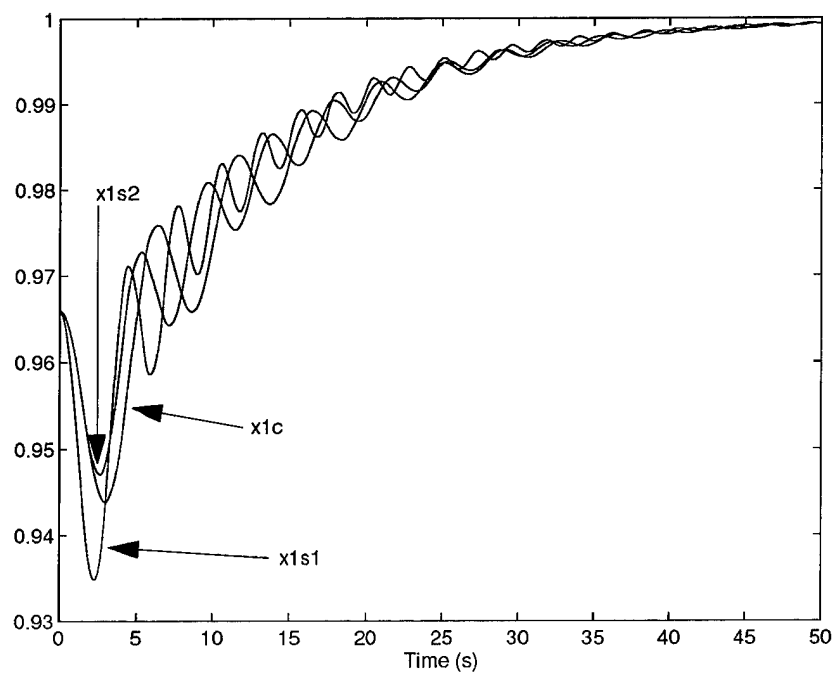


Figure 12.15: Prolate  $x_1$  Histories for 5-State Gyrostat

used, and the initial condition was  $(x_3, x_2, \mu_2, \nu) = (\sin(75^\circ), 0, 0, 0.2 \cos(75^\circ) - 1)$ , corresponding to the all-spun condition at an initial  $75^\circ$  cone angle. In Figures 12.16 and 12.17 we show the state time histories for the SDRE nonlinear regulator with  $H = H_1$  and  $H = H_2$ , respectively.

Note that for  $H = H_1$ ,  $\nu$  is driven to zero fairly quickly (within about fifty seconds), but the  $x_2$  and  $x_3$  states have large initial oscillations, while for  $H = H_2$ , the initial oscillations of  $x_2$  and  $x_3$  are comparatively smaller than for  $H = H_1$ , but the  $x$  states decay more slowly, and  $\nu$  takes much longer to go to zero. These behaviors are intuitively satisfying in that they reflect either the zero or nonzero penalty on integrals of deviations of the  $x_2$  and  $x_3$  states in the two cost functions, and the resulting emphasis on controlling either  $\nu$  only ( $H = H_1$ ) or a weighted combination of all the states ( $H = H_2$ ). Note that it is not apparent from Figure 12.17 that  $\nu$  actually does go to zero. We therefore show the continuation of the trajectory in this case in Figure 12.18, and it can be seen that  $\nu$  does indeed eventually go to zero, although it takes quite a while. This settling time may be decreased by increasing the relative weight on  $\nu$ , if desired. For comparison purposes we show state histories for the  $u = 0.01$  constant torque transverse spinup maneuver in Figure 12.19. Note that the large initial oscillations in the  $x_2$  and  $x_3$  states continue to occur longer in the constant torque case, and the decay rate appears slower than in either of the SDRE cases. In Figures 12.20 and 12.21 we show control and  $x_1$  histories for the three methods.

Note that, as expected, the  $H = H_2$  SDRE controller does the best job of minimizing deviations of  $x_1$  from the desired value of one, but uses a larger initial control value to accomplish this. Note also that the other two methods allow  $x_1$  to go negative for this simulation, but that does not prevent effective regulation of the states for these methods. It is really only the  $H = H_2$  SDRE controller that needs  $x_1 > 0$  because (recall that) at  $x_1 = 0$  we are no longer guaranteed a stabilizable four-state factorization. Thus, starting from larger initial cone angles than  $75^\circ$ , it may not be possible to use the  $H = H_2$  SDRE method since we are likely to lose stabilizability of the factorization when  $x_1$  goes negative. This problem technically also applies to the  $H = H_1$  SDRE controller, but can be avoided by solving only the scalar SDRE involving  $\nu$ , which is always controllable and observable if

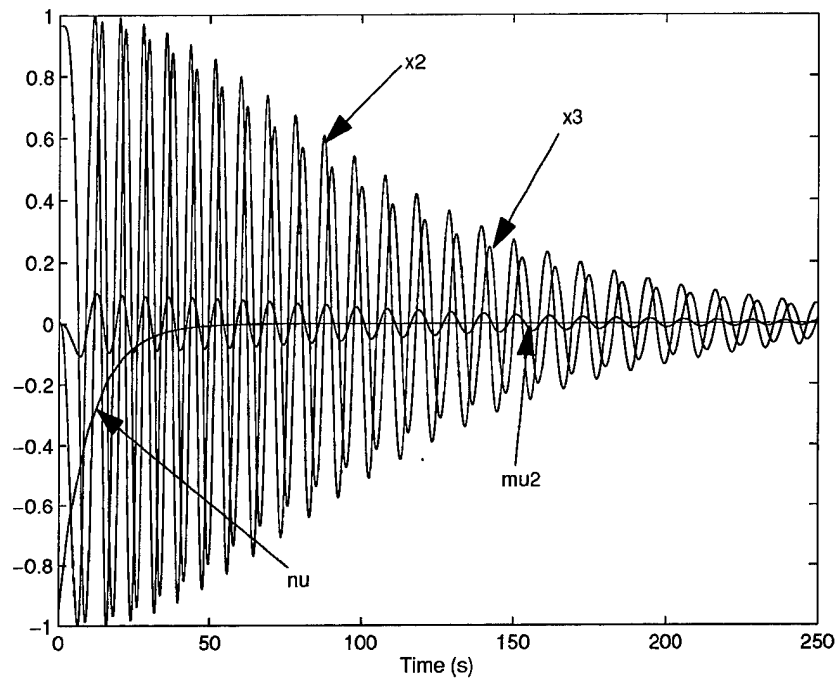


Figure 12.16: Transverse SDRE State Histories for 5-State Gyrostat ( $H = H_1$ )

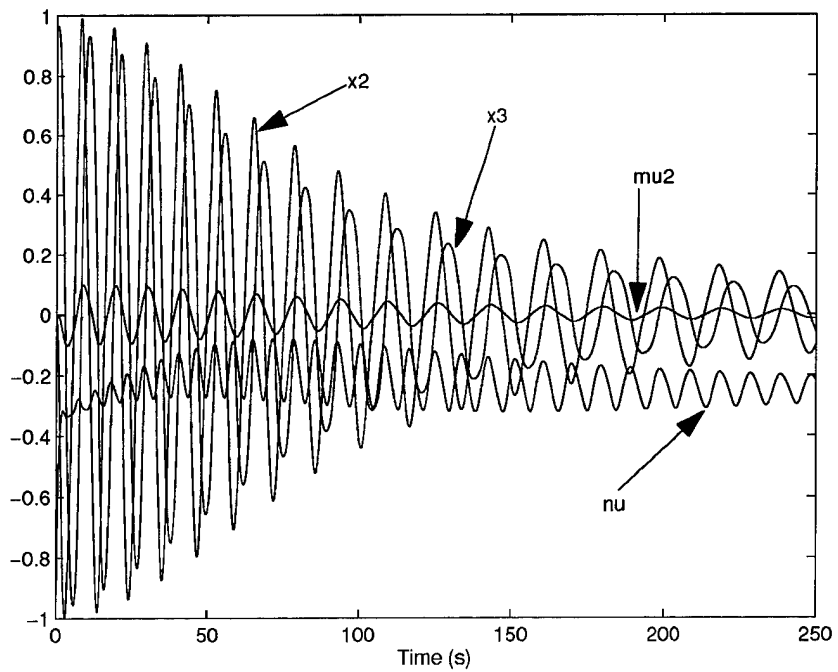


Figure 12.17: Transverse SDRE State Histories for 5-State Gyrostat ( $H = H_2$ )

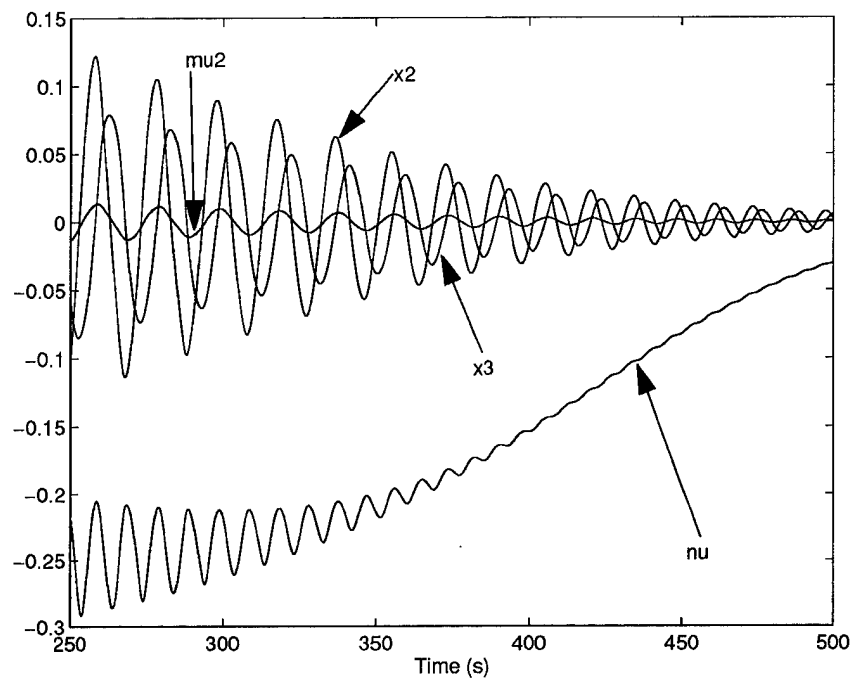


Figure 12.18: Transverse SDRE State Histories for 5-State Gyrostat ( $H = H_2$ )

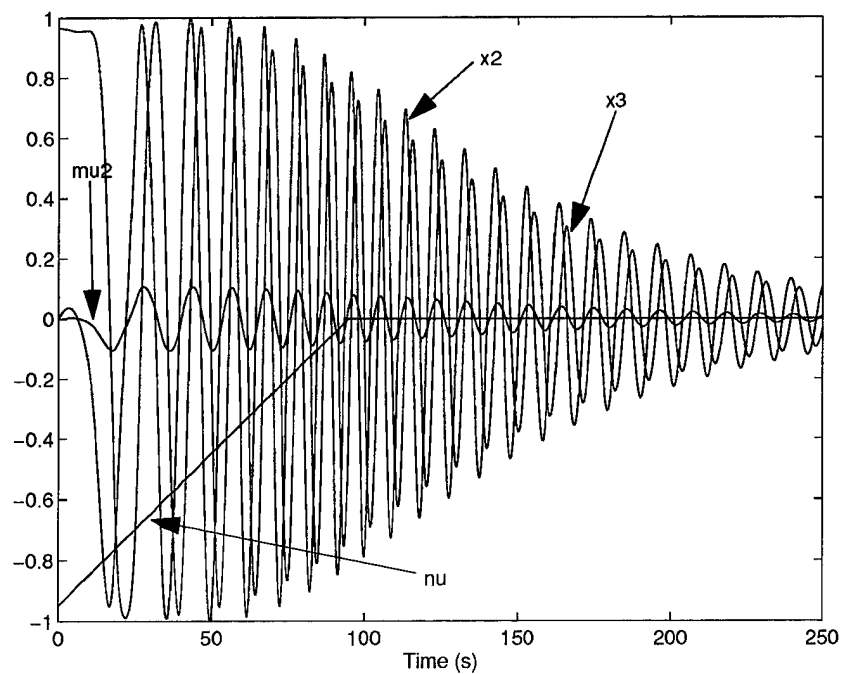


Figure 12.19: Transverse State Histories for 5-State Gyrostat ( $u = 0.01$ )

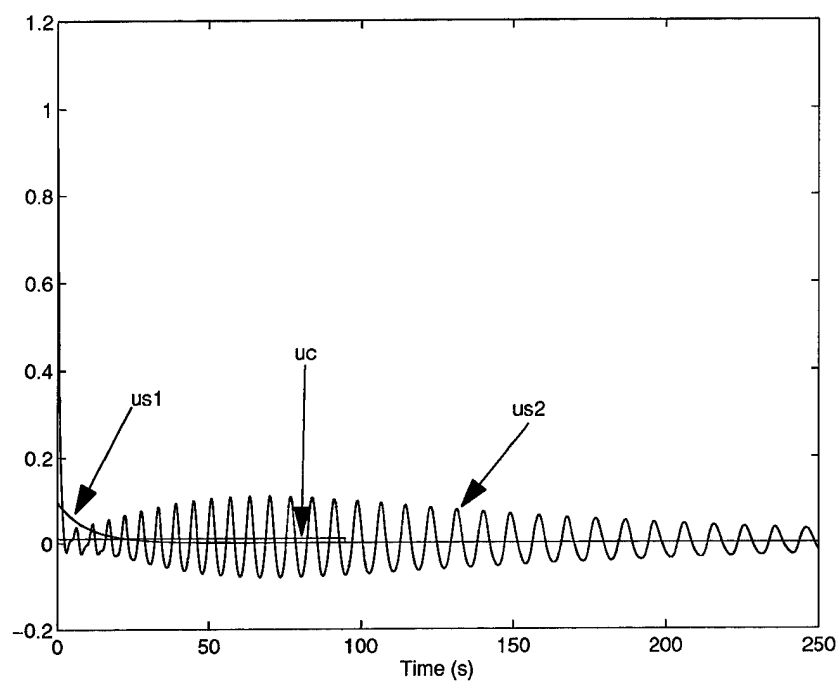


Figure 12.20: Transverse Control Histories for 5-State Gyrostat

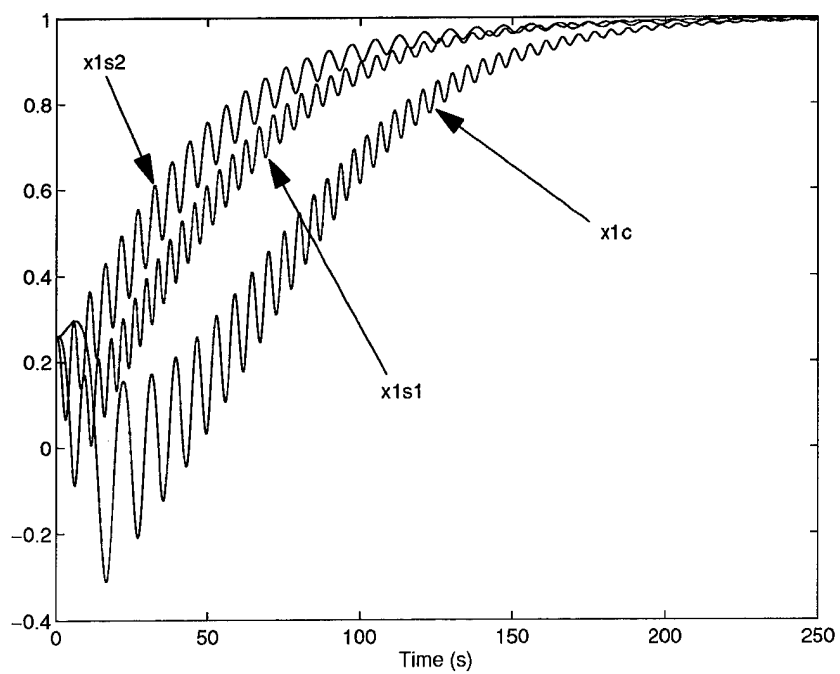


Figure 12.21: Transverse  $x_1$  Histories for 5-State Gyrostat

$\nu$  is penalized appropriately. For this simulation, however, the four-state SDRE was solved for the  $H = H_1$  case all along the trajectory without any difficulties.

### 12.7 SDRE Nonlinear $H_\infty$ Simulation Results

In this section we briefly discuss application of the SDRE nonlinear  $H_\infty$  control algorithm to the gyrostat example problem. Since we already know that no computational SDRE method can stabilize the original 4-state gyrostat model, our discussion here applies to the same model as used for the regulator: the 5-state gyrostat model with off-axis rotor added. Note that we actually have a 4-state design model, since we remove the  $x_1$  state. If we recall the three additional assumptions of Section 11.4 necessary to extend the sampled data SDRE nonlinear regulator theory to the nonlinear  $H_\infty$  case:

- i.  $G(x)$  is globally analytic with respect to  $x$
- ii. The pair  $\{A(x), K(x)\}$  is globally stabilizable
- iii. Stabilizing solutions to (11.51) exist everywhere along the trajectory

and we further invoke the assumption  $K \leq 0 \forall x$  sufficient to guarantee satisfaction of Assumption iii above, then it is easily seen that  $G(x)G^T(x)$  must be of the form

$$G(x)G^T(x) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g_4^2(x) \end{bmatrix} \quad (12.59)$$

and additionally we must have

$$g_4^2(x)/\gamma^2 - 1 \leq 0 \forall x \quad (12.60)$$

Thus, under these assumptions, the nonlinear  $H_\infty$  SDRE is equivalent to the nonlinear regulator SDRE with  $BR^{-1}B^T$  set equal to  $B(1 - g_4^2(x)/\gamma^2)B^T$  where  $0 \leq 1 - g_4^2(x)/\gamma^2 \leq 1$ , so that at each  $x$  the nonlinear  $H_\infty$  SDRE is equivalent to the regulator SDRE with increased control penalty. If

$g_4(x)$  is assumed constant, then the pointwise equivalence between the nonlinear  $H_\infty$  and scaled nonlinear regulator SDREs is global. However, note that since  $u = -R^{-1}B^TPx$  does not see the equivalent scaling effect on  $R$ , the nonlinear  $H_\infty$  control is not the same as the control from the scaled regulator problem. If the two were the same, we would expect to see reduced control effort for the nonlinear  $H_\infty$  problem, as per the effect of higher  $R$  in the regulator problem. However, since the effect of increased  $R$  in only the SDRE is to make the system appear less strongly controllable (it makes the magnitude of  $B$  appear smaller), we should expect higher gain solutions for the SDRE nonlinear  $H_\infty$  control problem. To illustrate this concept, we show comparative SDRE regulator and nonlinear  $H_\infty$  simulation results for the  $15^\circ$  initial cone angle oblate case with  $H = H_1$  in Figure 12.22. Note that only the control and  $\nu$  state histories are plotted, and the 'hi' and 'r' suffixes in the figure correspond to the  $H_\infty$  and regulator cases, respectively. The other state histories are not plotted because they are virtually indistinguishable for about the first thirty seconds, after which there is a slight phase difference in the oscillatory state histories. We have assumed

$$G(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ g_4 \end{bmatrix} \quad (12.61)$$

and  $g_4^2/\gamma^2 = 1/2$ . The effect of increased  $R$  in the SDRE is, as expected, higher initial control gain and thus somewhat faster regulation response. In fact, for this case the SDRE solutions are constant and are  $p_{44} = 0.1414$  and  $p_{44} = 0.1$ , respectively, for the nonlinear  $H_\infty$  and regulator cases, with all other  $p_{ij} = 0$ . This initial higher gain behavior in the nonlinear  $H_\infty$  setting carries over to the  $H = H_2$  case as well, as illustrated by Figure 12.23, in which we have plotted the two control histories for this case. To conclude, we note that the SDRE nonlinear  $H_\infty$  theory for this gyrostat example problem, while not adding much in the way of significant design flexibility due to the rank deficiency of  $B$  and corresponding strictness of assumptions made to guarantee existence of SDRE solutions, nevertheless is seen to yield stabilizing controllers under the appropriate additional assumptions.

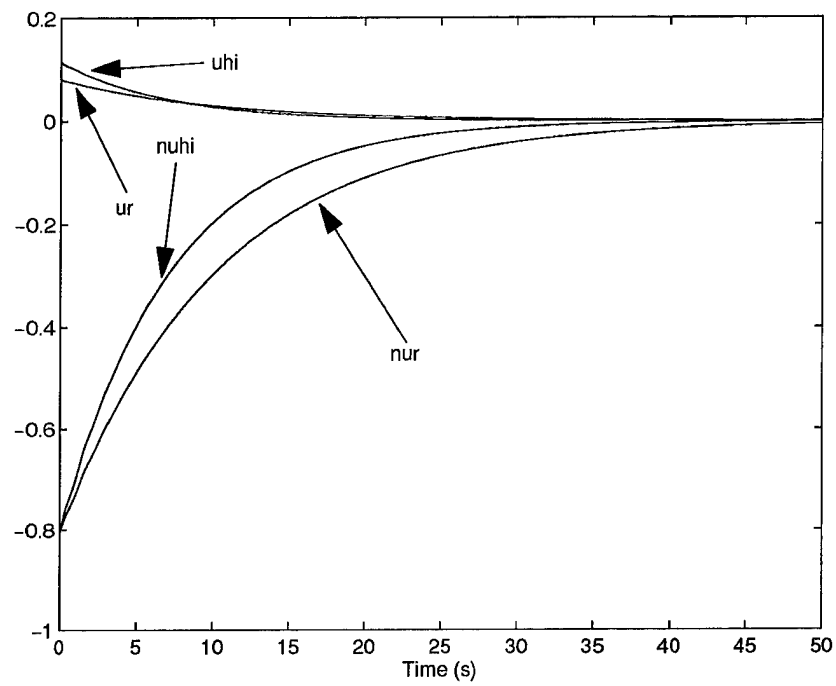


Figure 12.22: SDRE Nonlinear  $H_\infty$  and Regulator Histories for Oblate 5-State Gyrostat ( $H = H_1$ )

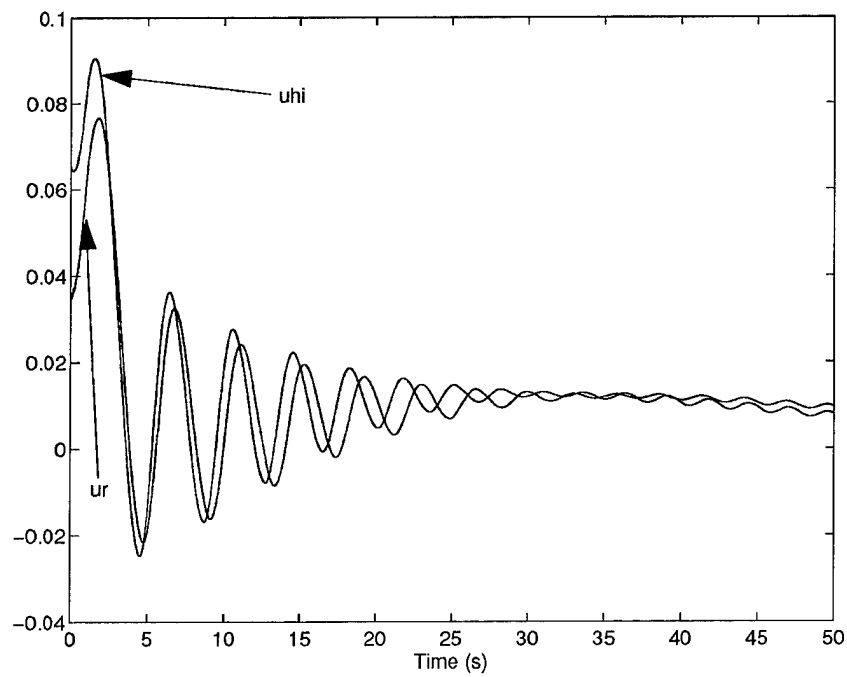


Figure 12.23: SDRE Nonlinear  $H_\infty$  and Regulator Controls for Oblate 5-State Gyrostat ( $H = H_2$ )

## 12.8 Comparison with Other Methods

In this section we briefly examine the applicability of feedback linearization and recursive backstepping to the gyrostat control problem, and compare their usefulness to that of the SDRE methods. We start with feedback linearization. Recall from Section 2.2 that, if we can solve the state space exact linearization problem, then we can turn the system into the form

$$\begin{aligned}\dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ \dot{z}_3 &= z_4 \\ \dot{z}_4 &= u\end{aligned}\tag{12.62}$$

which is a controllable linear system that is easily stabilized to the origin. To solve this problem, however, recall we need an output function for which (12.2) has relative degree four, and such an output exists if and only if the two controllability-like conditions of Theorem 2.2.3 are satisfied. Thus  $\Delta_3$  of (12.6) evaluated at  $(x_1, x_2, x_3, \mu) = (1, 0, 0, 1)$  must be full rank. This matrix is only rank one for this problem, even if we go to the reduced system after eliminating the  $x_1$  state. Thus, the state space exact linearization problem is not solvable for the baseline gyrostat. We are not really interested in an arbitrary change of coordinates to achieve relative degree four anyway, since we have two known outputs, namely  $x_1$  and  $\mu$ , which we desire to drive to selected values, and they are not related by a pure derivative relationship. Since in feedback linearization we can independently control only a number of outputs strictly less than or equal to the number of inputs, we are thus more interested in applying feedback linearization to the gyrostat with either  $x_1$  or  $\mu$  selected as the output to control. If we pick  $\mu$  as our controlled output, and we want to drive it to one, this is again accomplished by the change of variables  $\nu = \mu - 1$  and regulation of  $\nu$ . Since we have  $\dot{\nu} = u$ , feedback linearization in this case simply gives the linear control law  $u = -k\nu = -k(\mu - 1)$ , where  $k > 0$  is a user selectable constant that controls how quickly  $\nu$  is regulated. The theory for input-output linearization then requires that the zero dynamics of the remaining system be asymptotically stable

in order to yield asymptotic closed loop stability, which can be true for this problem only if we consider the gyrostat with off-axis rotor to provide damping. It is interesting to note that this same family of controls is achieved through the SDRE nonlinear regulator by taking  $k_1 = k_2 = k_3 = 0$  and  $k_4 = k$  in (12.54). Thus, the SDRE framework allows us to recover all controls that feedback linearization offers for this choice of controlled output, and both methods require the addition of the off-axis rotor to give asymptotic closed loop stability.

If, on the other hand, we choose  $x_1$  as the controlled output, it has relative degree two and so must be differentiated twice for the control to appear in the associated dynamics. We have to define  $\hat{x}_1 = x_1 - 1$  to be able to drive  $x_1$  to one, and we then get a system of the form

$$\begin{aligned}\dot{x}_2 &= a_2(x) \\ \dot{x}_3 &= a_3(x) \\ \dot{\mu} &= u \\ \dot{\hat{x}}_1 &= z_1 \\ \dot{z}_1 &= z_2 \\ \dot{z}_2 &= (i_2 - i_3)[a_1(i_3 x_3^2 - i_2 x_2^2) + 4a_2 a_3 + (x_2^2 - x_3^2)u]\end{aligned}\tag{12.63}$$

where the  $a_i$  are the right hand sides of the original dynamics equations with  $x_1 = \hat{x}_1 + 1$  substituted.

Thus, to stabilize the linearized subsystem of (12.63), we choose a control law of the form

$$u = \frac{1}{(i_2 - i_3)(x_2^2 - x_3^2)} \{-(i_2 - i_3)[a_1(i_3 x_3^2 - i_2 x_2^2) + 4a_2 a_3 + g^T z]\}\tag{12.64}$$

where  $g$  is a constant gain vector which stabilizes the characteristic equation

$$\lambda^3 - g_3 \lambda^2 - g_2 \lambda - g_1 = 0\tag{12.65}$$

We note that this control law is numerically ill-conditioned for small  $x_2$  and  $x_3$ , and undefined when  $x_2^2 = x_3^2$ . In addition, we still must have stability of the zero dynamics to get overall closed loop stability, and even if we obtain stability, we still cannot use this control law to drive  $\mu$  to one, since the control is already specified. If we add the off-axis rotor equation to (12.63), we get a feedback

linearizing control of the form

$$u = \frac{1}{(i_2 - i_3)(x_2^2 - x_3^2) + \alpha_2 \mu_2 x_2} \{-\psi(x) + g^T z\} \quad (12.66)$$

where  $\psi(x)$  is a complicated function of the states, so that the control is still ill-defined as the trajectory approaches the desired equilibrium point. Attempts to simulate this feedback linearization control law exhibited significant numerical instability. This control law, then, has several unattractive features which do not allow us to accomplish all that we desire for this problem. The conclusion to be reached from all this is that the SDRE nonlinear regulation method offers everything that feedback linearization offers for this problem, plus the additional design flexibility allowed by choosing to penalize multiple outputs and combinations of states in the SDRE controlled outputs.

In applying recursive backstepping to the four-state gyrostator, we see that the system is not in the assumed ideal strict-feedback form, but that by considering  $x_1$ , either  $x_2$  or  $x_3$ , and  $\mu$ , we have a system in strict-feedback form plus a one-state zero dynamics subsystem. Thus, we consider recursive backstepping applied to

$$\dot{x}_1 = cx_2x_3 \quad (12.67)$$

$$\dot{x}_2 = i_3x_1x_3 - x_3\mu \quad (12.68)$$

$$\dot{\mu} = u \quad (12.69)$$

where  $c = i_2 - i_3 > 0$ , and we treat  $x_3$  as a parameter with known dynamics, and we assume it can be measured perfectly in real time. To apply recursive backstepping, we need to select a stabilizing function, or desired value, for  $x_2$  in (12.67), but first we need to shift the origin to be our desired final equilibrium point. Thus, let  $z_1 = x_1 - 1$  and  $\nu = \mu - 1$  as before, to get

$$\dot{z}_1 = cx_2x_3 \quad (12.70)$$

$$\dot{x}_2 = i_3(z_1 + 1)x_3 - x_3 - x_3\nu \quad (12.71)$$

$$\dot{\nu} = u \quad (12.72)$$

Choosing the stabilizing function

$$\hat{x}_2 = -x_3z_1 \quad (12.73)$$

would give  $\dot{z}_1 = -x_3^2 z_1$  if achieved perfectly, which is stable for all  $x_3$  and asymptotically stable if  $x_3 \neq 0$ . We thus have the error function

$$e_2 = x_2 + x_3 z_1 \quad (12.74)$$

so that the dynamics become

$$\dot{z}_1 = c e_2 x_3 - k x_3^2 z_1 \equiv f_1 \quad (12.75)$$

$$\dot{e}_2 = i_3(z_1 + 1)x_3 - x_3 - x_3 \nu + x_3 f_1 + z_1[\nu + 1 - i_2(z_1 + 1)][e_2 - x_3 z_1] \equiv f_2 \quad (12.76)$$

$$\dot{\nu} = u \quad (12.77)$$

Proposing the Lyapunov function  $V = 0.5(z_1^2 + e_2^2 + \nu^2)$  we find

$$\dot{V} = z_1 f_1 + e_2 f_2 + \nu u \quad (12.78)$$

and we see that, if we have  $\dot{V} < 0$  globally, we drive  $z_1$ ,  $e_2$ , and  $\nu$  to zero. This is equivalent to driving  $x_1$  to one, and also by (12.74) to driving  $x_2$  to zero, which by the constraint  $x_1^2 + x_2^2 + x_3^2 = 1$  gives  $x_3 = 0$ , so that our design objective would be achieved. However, when  $\nu = 0$ , by (12.78) we cannot use the control to make  $\dot{V}$  negative, and expansion of the terms in (12.78) shows that  $\dot{V}$  is not negative definite for  $\nu = 0$ . Choosing to use  $x_3$  instead of  $x_2$  in the above development does not change this fact, so that this recursive backstepping approach will not allow us to achieve our design objectives.

Alternate approaches to applying recursive backstepping to this problem could be tried along the lines of the approaches used in the feedback linearization discussion above. Theorem 2.3.2 would allow us to choose a control of the form  $u = -k(\nu)\nu$ , where  $k$  is a globally positive definite function of  $\nu$ , provided we knew a Lyapunov function for the remaining zero dynamics subsystem with guaranteed negative derivative. Without the off-axis rotor, we know such a Lyapunov function does not exist since the system is only stable, and not asymptotically stable. We might be able to find such a function for the five-state gyrostat, however, in which case we recover the  $H = H_1$  SDRE approach or equivalently, the  $\nu$ -controlled variable feedback linearization approach, if  $k(\nu)$  is

chosen a positive constant. In addition, both the SDRE and recursive backstepping approaches can generate nonconstant positive definite  $k(\nu)$ , simply by choosing the appropriate control in recursive backstepping, and by choosing nonconstant, positive definite  $k_4(\nu)$  in the  $H$  factorization (12.54) for the SDRE nonlinear regulator.

Finally, recursive backstepping could be applied to the  $x_1$  subsystem differentiated twice, as in the second feedback linearization approach. This approach, although it allows more flexibility than feedback linearization, is nevertheless like the feedback linearization approach in that it leads to an ill-defined control when trajectories approach the origin, and also leaves  $\nu$  uncontrolled, and thus is not particularly useful.

As was mentioned in Chapter 2.3, recursive backstepping is quite user-dependent, and thus no general claims of its applicability or lack thereof to this problem can be made. By this, we mean another designer might be able to apply the method to this problem with more success than we have. What may be concluded, however, is that in this section we have tried the most obvious applications of the method, and have again found that SDRE offers just as much as recursive backstepping does for these approaches in solving this problem.

## 12.9 Conclusion

In this chapter we have successfully applied the sampled data SDRE nonlinear regulation algorithm to perform spinup maneuvers of an axial gyrost, and we investigated the utility of other control algorithms in performing the same maneuvers. First, we saw that none of the methods could asymptotically stabilize the system without the addition of the off-axis damping rotor, and so our system configuration was modified to include its presence. For the five-state gyrost, the SDRE controllers were compared to a constant torque control that is commonly used to perform the desired maneuvers, and significant improvement in controlling  $x_1$  over the constant torque control was achieved in the prolate and transverse spinup maneuvers. An observed strength of the SDRE nonlinear regulator is that it offered significant design flexibility, while also stabilizing the system to the desired

equilibrium point. This flexibility was established by finding factorizations of the dynamics which guaranteed well-posedness of the control in a hemisphere of the momentum sphere, and penalizing the control and different sets of states. By penalizing only  $\nu$  with a constant weight, a stabilizing linear control law was obtained, and it was shown that this family of control laws is the same as that obtained by applying feedback linearization to the gyrostat problem. By allowing nonconstant weights on  $\nu$  only, an intuitive set of achievable controls from recursive backstepping was recovered. The SDRE algorithm, however, also allows the  $x_2$  and  $x_3$  states to be influenced, in that by penalizing these states, the control can be made to actively reduce their deviations from zero over time. SDRE nonlinear regulation thus provides a systematic way of achieving blended performance in terms of driving both  $x_1$  and  $\mu$  to the desired operating points, and this is a capability that none of the other methods offer.

### *XIII. Conclusions and Recommendations for Further Research*

In this chapter we include a brief summary of this research, highlighting the specific contributions made. We draw some conclusions based on the results obtained, and conclude with recommendations for further research.

#### *13.1 Summary of Conducted Research*

In this dissertation, we have conducted a fairly thorough survey of modern nonlinear control design techniques, identifying the current state of the art in the most common methods, and thus, highlighting potential areas for research. Following this survey given in Chapter 1, we directed our focus to the new nonlinear control design techniques based on state-dependent Riccati equations (SDREs). This choice of focus was motivated by three major factors: recent success in practical application of the SDRE methods; a documented history in the literature of applying linear control techniques to pointwise linear factorizations of nonlinear systems; and a significant lack of theoretical justification for use of these methods, or for logical choices of factorizations. To obtain a basis for comparison with the SDRE based methods, we selected three common design methods, namely feedback linearization, recursive backstepping, and nonlinear  $H_\infty$  control, and we included the basic theoretical components for control design using each technique in Chapter 2. We have limited our scope to the case in which full state feedback is available, and we have focused our attention mainly on achieving closed loop stability of nonlinear systems, but we have also been interested in performance issues using these methods. To facilitate stability analysis, we have included a number of tools, including the *Principle of Stability in the First Approximation* [68], Lyapunov theory [40], Lasalle's Invariance Principle [44], and center manifold theory [10].

To enable comparison between the four nonlinear control methods given above and to illustrate promising areas of research, we applied the methods to a motivational two-state example problem in Chapter 3. From this experience we were able to make several observations. It appears that feedback

linearization is the most restrictive method in terms of the conditions which a system must satisfy for its use, and we were also able to see that simply canceling nonlinearities, as feedback linearization does, is not always the best design approach. Next, we saw that recursive backstepping is an effective tool for achieving stabilization, but other performance objectives are not easily incorporated. Although the example problem did not illustrate this, a potential issue for recursive backstepping is getting systems into the required strict-feedback form, and what to do when systems are not transformable into such a form. This issue was indeed a factor for the gyrostat example problem. Nonlinear  $H_\infty$  control theory, as opposed to recursive backstepping, was seen to provide a nice general framework for posing control problems which allow easy incorporation of performance as well as closed loop stability objectives. While we were able to manipulate a proposed polynomial form solution to give a satisfactory local answer for the example problem of Chapter 3, general methods of solving the resulting Hamilton-Jacobi inequalities (Issacs equations) were nevertheless needed. The SDRE method was seen to offer promise in this regard; however, a number of issues having to do with using a storage function as a Lyapunov function needed to be worked out. The SDRE nonlinear regulator theory as presented by Cloutier, *et al.* [13] was also promising, yet this method was seen to have the same challenge of guaranteeing closed loop stability, and raised additional questions about choices of SDC parametrizations, and the impact of controllability/observability tests based on the standard linear theory.

Based on the results of Chapter 3 and a detailed study of the existing theory for use of SDRE methods, a number of design issues were addressed in Chapter 4. To enable subsequent analysis, formulas for partial derivatives of vector matrix products with  $x$  dependency were given in Section 4.1. The implications of convexity to sufficiency of local optimal solutions and global optima were then addressed, and a sufficient condition for convexity of the optimal regulator cost function was given, using the above-mentioned vector matrix product gradient results. In the next several sections, issues concerning existence of optimal cost functions and their use as system Lyapunov functions were clarified. In particular, symmetry issues of the SDRE solutions were addressed, and

a simple form sufficient condition for satisfying the first-order necessary conditions for optimality was derived. It was seen that, although local solutions satisfying these first-order necessary conditions were obtainable for simple systems, extending the region of validity of the results, especially for high dimensional systems, was prohibitive in terms of development and implementation costs. Thus, the decision was made to focus on suboptimal solutions, and to seek means of guaranteeing stability other than through the use of optimal cost functions as Lyapunov functions. Also, analytic solution methods for SDREs were discarded, based on effort expended in failed attempts to solve simple problems, in favor of numerical solutions using standard software. Finally, a comparison between HJIs and HJIEs was made, and a necessary condition for solvability of HJIs and HJB equations was given, introducing the notions of nonlinear controllability and stabilizability.

In the next several chapters, we focused on the need to provide stability guarantees for closed loop systems using suboptimal nonlinear SDRE regulators. In Chapter 5, necessary and sufficient conditions for locally asymptotically stabilizing analytic solutions for scalar analytic systems were derived and presented. In Chapter 6, the general nonequivalence relationship between true nonlinear controllability and linear controllability of SDC factorizations of nonlinear systems was developed, and the impact of nonlinear stabilizability on successful application of SDRE methods was assessed. Particular cases where both types of controllability hold were derived, and one of these corresponding to full rank, constant  $B$  matrices motivated the global asymptotic stability proof for SDRE regulators using positive definite  $Q$  matrices given in Chapter 7. In the next two chapters, sampled data implementations of continuous SDRE regulators were considered, and conditions guaranteeing semiglobal asymptotic closed loop stability were presented and proven. Key elements of the analysis in these two chapters were state transition matrix representations for nonlinear, discrete time systems; definition of an appropriate Lyapunov function; and positive semidefiniteness of the state weighting matrix  $Q$  leading to invariant set theory in the SDRE framework, as needed for application of Lasalle's Invariance Principle. The theory of Chapters 8 and 9 greatly clarified a number of issues regarding stability of SDRE regulators and appropriate selection of SDC factorizations. Then, in

Chapter 10, a line of inquiry leading to semiglobal exponential stability of SDRE regulators was investigated, showing the potential for such a characteristic to exist. To conclude our theoretical development, in Chapter 11 we showed that all the stability analysis performed for the SDRE regulator in Chapters 5 through 10 could be extended to the SDRE nonlinear  $H_\infty$  setting, provided suitable additional assumptions are made. We also gave or outlined modified proofs for many of the stability theorems in the  $H_\infty$  case, showing the impact of the needed assumptions.

In Chapter 12, a nontrivial axial gyrostat example problem was investigated, illustrating the SDRE stability and factorization theory developed in the previous chapters. Comparison with feedback linearization, recursive backstepping, and constant torque controllers was also performed, and the SDRE regulator was seen to allow great design flexibility and superior performance in stabilizing the system to the desired equilibrium point. It was also briefly discussed how the assumptions guaranteeing existence of SDRE solutions in the nonlinear  $H_\infty$  case limit the applicability of the method to the gyrostat problem, in effect yielding solutions equivalent to SDRE regulators with larger magnitude control penalties. Since the apparent increase in control penalty  $R$  was seen only in the SDRE solution, these nonlinear  $H_\infty$  solutions were seen to produce larger initial magnitude controls than their nonlinear regulator counterparts.

### *13.2 Conclusions*

In light of the above discussion, we now conclude by identifying some of the specific contributions this research has made, and attempting to assess their impact. Through this dissertation, we have made substantial progress in developing a coherent and rigorous theory for use of the SDRE methods. We have established sufficient conditions for the SDRE regulator to be locally and globally optimal and stable, including determining factorization-dependent conditions for positive definiteness of the corresponding Lyapunov function. Making use of the simplified necessary condition for optimality we derived, we have proposed and validated a numerical local solution scheme to the optimal problem for second-order systems. We have expressly chosen not to investigate globally optimal solution

algorithms for general systems, because we have seen that the development and implementation costs for such an algorithm are quite likely to exceed the potential gain to be had. We have established a necessary condition for solving general HJIEs and HBJ equations involving nonlinear stabilizability, and we have further made significant progress in finding solutions to these equations using the SDRE methods. We have examined in detail two proposed solution forms to the nonlinear  $H_\infty$  HJI via SDRE, concluding that symmetric SDRE solutions only need be considered, in part because we have observed that analytical solutions will not in general be feasible. Under conditions of symmetry, we have also demonstrated a clear relationship between the two solution methods, indicating one of the methods to be more useful, since the other then becomes just a special case of the first, solvable only for linear time invariant systems. We have also investigated solution of strict HJIs versus HJIEs, showing that solving a strict HJI is equivalent to solving an HJIE with a positive definite penalized output norm. Thus, solving strict HJIs is a simpler way to guarantee closed loop stability, but it is in general a harder problem, which may restrict solvability for a given problem.

In the latter chapters of this dissertation, we have made significant strides in determining conditions which guarantee closed loop stability of both SDRE regulators and SDRE nonlinear  $H_\infty$  controllers, in both continuous time and sampled data formats. For scalar analytic systems, we have established necessary and sufficient conditions for analytic locally stabilizing solutions, showing that stabilizability/detectability of the regulator SDC factorization is sufficient, but not at all necessary for local stability (as also observed earlier in the two-state example problem of Chapter 3). We examined true and factored controllability of nonlinear and SDC systems, establishing a local equivalence relationship between the two, and other special cases when equivalence holds, although the two concepts are, in general, different. Using this knowledge, the matrix measure and other researchers' work [49] on bounds of solutions to algebraic Riccati equations, we were able to prove global asymptotic stability of continuous time systems with full rank, constant  $B$  matrices, and positive definite  $Q$  matrices, assuming some slight additional conditions were satisfied. The sufficiency of having stabilizable and detectable factorizations for well-posedness of the numerically computed

SDRE regulator was established, based on existence of stabilizing solutions to the SDRE. By focusing on numerical, stabilizing solutions to AREs, we forego taking advantage of the observed ability of SDRE methods to stabilize systems locally that do not possess stabilizable and detectable factorizations, however. It was also discovered that stabilizability/detectability of the SDC parametrizations gives analytic SDRE solutions, a fact which was used to help prove asymptotic stability of sampled data SDRE regulators. The effect of stabilizability and detectability in factorizations on closed loop stability was also identified, based on pointwise decompositions of nonlinear systems and invariant set theory. In particular, we saw that states in the null space of  $H(x)$  do not receive any control action, so that invariant sets in the null space of  $H$  must be asymptotically stable in order to obtain closed loop stability. Detectability of the factorization does not, in general, guarantee this, so that additional analysis may be necessary. Choosing  $H$  to be globally nonsingular so that  $Q$  is positive definite is then well motivated to achieve a negative Lyapunov function derivative, but may not lead to a stable closed loop system when the system is not globally nonlinearly stabilizable, as examples illustrated. Factorized stabilizability, on the other hand, was seen to guarantee that convergent closed loop solutions will not be attracted to the uncontrollable space, so that the above analysis based on observability concepts plays the determining role in guaranteeing closed loop stability. We then saw that, if one can choose the  $Q$  matrix function to give a diagonalizable and sufficiently stable closed loop dynamics matrix function,  $F$ , then exponential stability of the sampled data SDRE regulator may be obtained. Finally, by making some additional assumptions on the disturbance to the system dynamics, we saw that the above arguments can be successfully extended to the SDRE nonlinear  $H_\infty$  case.

We can also draw several conclusions from application of the SDRE nonlinear regulator to the gyrostat problem. We saw that successful application of the method hinges on being able to find a stabilizable and detectable SDC factorization for the system, and that finding such a parametrization can be a nontrivial exercise for larger dimension systems. Intuitively speaking, attempting to maximize the controllable and observable spaces of the factorizations in a linear sense is a good rule

of thumb for selecting parametrizations, since pointwise control effort can be directly linked to these issues. We also saw a potential limitation of the numerical implementation of SDRE methods, in that they revert to standard linear design techniques acting on the linearized dynamics in a neighborhood of the origin, and thus may or may not stabilize systems that a purely nonlinear controller could. As seen in Chapter 3, controls based on analytic solutions to the SDRE can potentially avoid this limitation. The simulation results verified the Lyapunov stability analysis given in this dissertation in that, given nonlinear controllability, we obtained convergent SDRE solutions, so that trajectories went to the equilibrium point at the origin. This result was expected, based on the fact that we had stabilizable and detectable factorizations, and sufficiently rapid sampling. Although the effect of various sampling rates was not illustrated for this problem, it was observed in other simulations conducted during this research that decreasing the sampling rate beyond a critical level can indeed lead to instability. In addition to providing the desired closed loop asymptotic stability for this problem, the SDRE method allowed significant design flexibility through the option of penalizing various combinations of states. In fact, it was seen that the SDRE technique could produce all the controllers one would reasonably expect from applying feedback linearization and recursive backstepping to this problem, plus additional controllers. This systematic design flexibility is therefore judged to be one important advantage the SDRE method offers. Finally, this flexibility was key to the SDRE nonlinear regulator being able to achieve better cone angle regulation performance than the commonly-used constant torque maneuver.

In conclusion, this research has provided theoretical justification and demonstration of applicability of SDRE-based control techniques. Stability analysis and factorization selection are two major areas where we have contributed, although other contributions have also been made. We have been able to make an objective evaluation and comparison of available nonlinear control methods with the SDRE methods, and have seen that SDRE offers a number of design advantages, including relative ease of application if appropriate factorizations are available.

### 13.3 Recommendations for Further Research

There are a number of promising areas for further research which have become obvious as this research has progressed. We give a brief list of such below.

- Develop coherent and rigorous theory for output feedback SDRE nonlinear regulation and  $H_\infty$  control
- Extend the theoretical analysis of SDRE methods to include neutrally stabilizing solutions to SDREs, thereby not requiring stabilizable and detectable factorizations
- Fully develop the theory for exponential stability of SDRE controllers
- Prove that boundedness, as opposed to convergence, of SDRE solutions is sufficient for semiglobal asymptotic stability of observable trajectories
- Extend the global asymptotic stability results for full rank, constant  $B$  matrices to stability of systems having  $B$  matrix functions with some constant submatrix
- Find a tight bound on the nonoptimality of the SDRE suboptimal nonlinear regulator
- Establish a theoretical proof of boundedness of SDRE solutions given full nonlinear controllability of the open loop system
- Analyze the global induced  $L_2$  gain properties of the SDRE nonlinear  $H_\infty$  control algorithm
- Treat the full-scale sampled data regulator problem for the equivalent discrete time system, as opposed to the sampled data implementation of the continuous control law

It is hoped that these areas of research will be effectively pursued, because complete theoretical developments of the above topics would greatly enhance the utility of the SDRE methods.

## Bibliography

- [1] T. Apostol. *Mathematical Analysis* (Second Edition). Reading, MA: Addison-Wesley, 1974.
- [2] Z. Artstein. "Stabilization with Relaxed Controls," *Nonlinear Analysis, TMA*-7:1163-1173 (1983).
- [3] J. Ball, J. W. Helton, and M. L. Walker. " $H_\infty$  Control for Nonlinear Systems via Output Feedback," *IEEE Transactions on Automatic Control*, AC-38:546-559 (1993).
- [4] J. A. Ball and J. W. Helton. " $H_\infty$  Optimal Control for Nonlinear Plants: Connection with Differential Games." *Proceedings of 28th Conference on Decision and Control*. 956-962. December 1989.
- [5] G. Becker. *Quadratic Stability and Performance of Linear Parameter Dependent Systems*. PhD dissertation, University of California, Berkeley, CA, 1993.
- [6] R. R. Bitmead and B. D. O. Anderson. "Lyapunov Techniques for the Exponential Stability of Linear Difference Equations with Random Coefficients," *IEEE Transactions on Automatic Control*, 25(4):782-787 (1980).
- [7] S. Bittanti, A. J. Laub, and J. C. Willems, editors. *The Riccati Equation*. Berlin: Springer-Verlag, 1991.
- [8] A. E. Bryson and Y.-C. Ho. *Applied Optimal Control*. Hemisphere, 1975.
- [9] J. A. Burghart. "A Technique for Supoptimal Control of Nonlinear Systems," *IEEE Transactions on Automatic Control*, 14(5):530-533 (1969).
- [10] J. Carr. *Applications of Center Manifold Theory*. Applied Mathematical Sciences, NY: Springer-Verlag, 1981.
- [11] H. D. Chiang, M. Hirsch, and F. Wu. "Stability Regions of Nonlinear Autonomous Dynamical Systems," *IEEE Transactions on Automatic Control*, 33(1):16-26 (1988).
- [12] J. Cloutier, "1995 Annual AFOSR Review, Progress at WL/Eglin AFB, WPAFB, OH," November 1995.
- [13] J. R. Cloutier, C. N. D'Souza, and C. P. Mracek. "Nonlinear Regulation and Nonlinear  $H_\infty$  Control Via the State-Dependent Riccati Equation Technique: Part 1, Theory." *Proceedings of the First International Conference on Nonlinear Problems in Aviation and Aerospace*. May 1996.
- [14] J. R. Cloutier, C. N. D'Souza, and C. P. Mracek. "Nonlinear Regulation and Nonlinear  $H_\infty$  Control Via the State-Dependent Riccati Equation Technique: Part 2, Examples." *Proceedings of the First International Conference on Nonlinear Problems in Aviation and Aerospace*. May 1996.
- [15] M. Dalsmo and O. Egeland. " $H_\infty$  Control of Nonlinear Passive Systems by Output Feedback." *Proceedings of 34th Conference on Decision and Control*. 351-352. December 1995.
- [16] M. Dalsmo and O. Egeland. "State Feedback  $H_\infty$  Control of a Rigid Spacecraft." *Proceedings of 34th Conference on Decision and Control*. 3968-3973. December 1995.
- [17] J. D'Azzo and C. Houpis. *Linear Control System Analysis and Design: Conventional and Modern* (Third Edition). New York: McGraw-Hill, 1988.

- [18] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis. "State-Space Solutions to Standard  $H_2$  and  $H_\infty$  Control Problems," *IEEE Transactions on Automatic Control*, *AC-34*(8):831–847 (August 1989).
- [19] D. Ehrler and S. R. Vadali. "Examination of the Optimal Nonlinear Regulator Problem." *Proceedings of the AIAA Guidance, Navigation, and Control Conference*. August 1988.
- [20] S. N. Elaydi. *An Introduction to Difference Equations*. NY: Springer-Verlag, 1996.
- [21] B. Friedland. *Advanced Control System Design*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [22] C. D. Hall. "Resonance Capture in Axial Gyrostats," *The Journal of the Astronautical Sciences*, *43*(2):127–138 (1995).
- [23] C. D. Hall. "Spinup Dynamics of Gyrostats," *Journal of Guidance, Control, and Dynamics*, *18*(5):1177–1183 (1995).
- [24] C. D. Hall. "Momentum Transfer in Two-Rotor Gyrostats," *Journal of Guidance, Control, and Dynamics*, *19*(5):1157–1161 (1996).
- [25] C. D. Hall and R. H. Rand. "Spinup Dynamics of Axial Dual-Spin Spacecraft," *Journal of Guidance, Control, and Dynamics*, *17*(1):30–37 (1994).
- [26] R. Hermann and A. J. Krener. "Nonlinear Controllability and Observability," *IEEE Transactions on Automatic Control*, *22*:728–740 (1977).
- [27] D. Hill and H. Moyland. "The Stability of Nonlinear Dissipative Systems," *IEEE Transactions on Automatic Control*, *AC-21*:708–711 (1976).
- [28] K. Hoffman and R. Kunze. *Linear Algebra* (Second Edition). Englewood Cliffs, NJ: Prentice-Hall, 1971.
- [29] J. Huang and C.-F. Lin. "Numerical Approach to Computing Nonlinear  $H_\infty$  Control Laws," *Journal of Guidance, Control, and Dynamics*, *18*(5):989–994 (Mar 1995).
- [30] Y. Huang and W. Lu. "Nonlinear Optimal Control: Alternatives to Hamilton-Jacobi Equation." *Proceedings of 35th Conference on Decision and Control*. 3942–3947. December 1996.
- [31] J. Imura, H. Maeda, T. Sugie, and T. Yoshikawa. "Robust stabilization of nonlinear systems by  $H_\infty$  state feedback," *Systems and Control Letters*, *24*:103–114 (1995).
- [32] J. Imura, T. Sugie, and T. Yoshikawa. "A Hamilton-Jacobi Inequality Approach to Nonlinear Strict  $H_\infty$  Control and Characterization of Some Gain Conditions." *Proceedings of 34th Conference on Decision and Control*. 4233–4238. December 1995.
- [33] A. Isidori. *Nonlinear Control Systems* (Third Edition). Berlin: Springer-Verlag, 1995.
- [34] A. Isidori and W. Kang. " $H_\infty$  Control via Measurement Feedback for General Nonlinear Systems," *IEEE Transactions on Automatic Control*, *AC-40*(3):466–472 (Mar 1995).
- [35] M. R. James and S. Yuliar. "Numerical Approximation of the  $H_\infty$  Norm for Nonlinear Systems," *Automatica*, *31*(8):1075–1086 (1995).
- [36] R. E. Kalman and J. E. Bertram. "Control System Analysis and Design by the Second Method of Lyapunov," *Journal of Basic Engineering*, *82D*(2):371–400 (1960).
- [37] W. Kang. "Nonlinear  $H_\infty$  Control and Its Application to Rigid Spacecraft," *IEEE Transactions on Automatic Control*, *40*(7):1281–1285 (1995).

- [38] T. Kato. *Perturbation Theory for Linear Operators*. New York: Springer-Verlag, 1966.
- [39] H. K. Khalil. *Nonlinear Systems* (Second Edition). Upper Saddle River, NJ: Prentice-Hall, 1996.
- [40] M. Krstic, I. Kanellakopoulos, and P. Kokotovic. *Nonlinear and Adaptive Control Design*. New York, NY: John Wiley and Sons, 1995.
- [41] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems*. NY: Wiley Interscience, 1972.
- [42] P. Lancaster and M. Tismenetsky. *The Theory of Matrices* (Second Edition). Computer Science and Applied Mathematics, Orlando: Academic Press, 1985.
- [43] S. H. Lane and R. F. Stengel. "Flight Control Design Using Nonlinear Inverse Dynamics," *Automatica*, 24(4):471-484 (1988).
- [44] J. P. LaSalle. "Stability Theory for Ordinary Differential Equations," *Journal of Differential Equations*, 4:57-65 (1968).
- [45] W. M. Lu and J. C. Doyle. " $H_\infty$  control of Nonlinear Systems: a Convex Characterization," *IEEE Transactions on Automatic Control*, 40(9):1668-1675 (1995).
- [46] D. G. Luenberger. *Optimization by Vector Space Methods*. NY: John Wiley & Sons, Inc, 1969.
- [47] D. L. Lukes. "Optimal Regulation of Nonlinear Dynamical Systems," *SIAM Journal of Control*, 7:75-100 (1969).
- [48] W. McEneaney and K. Mease. "Nonlinear  $H_\infty$  Control of Aerospace Plane Ascent." *Proceedings of 34th Conference on Decision and Control*. 3994-3995. December 1995.
- [49] T. Mori and A. Derese. "A Brief Summary of the Bounds on the Solution of the Algebraic Matrix Equations in Control Theory," *International Journal of Control*, 39(2):247-256 (1984).
- [50] T. Mori, N. Fukuma, and M. Kuwahara. "Explicit Solution and Eigenvalue Bounds in the Lyapunov Matrix Equation," *IEEE Transactions on Automatic Control*, 31(7):656-658 (1986).
- [51] A. W. Naylor and G. W. Sell. *Linear Operator Theory in Engineering and Science*. Applied Mathematical Sciences, NY: Springer-Verlag, 1982.
- [52] H. Nijmeijer and A. J. van der Schaft. *Nonlinear Dynamical Control Systems*. Berlin: Springer-Verlag, 1990.
- [53] D. Parrish. *Applications of Nonlinear Control Using the State-Dependent Riccati Equation*. MS thesis, Air Force Institute of Technology, WPAFB, OH, 1995.
- [54] J. D. Pearson. "Approximation Methods in Optimal Control," *Journal of Electronics and Control*, 13:453-465 (1962).
- [55] Z. Qu, J. R. Cloutier, and C. P. Mracek. "A New Sub-optimal Nonlinear Control Design Technique-SDARE." *Proceedings of the 13th IFAC World Congress*. 365-370. July 1996.
- [56] D. B. Ridgely. *A Nonconservative Solution to the General Mixed  $H_2/H_\infty$  Optimization Problem*. PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1992.
- [57] D. B. Ridgely and S. S. Banda. *Introduction to Robust Multivariable Control*. AFWAL-TR-85-3102, USAF, 1986.

- [58] L. Rodman. "On Extremal Solutions of the Algebraic Riccati Equation." *Algebraic and Geometric Methods in Linear Systems Theory* Lectures in Applied Mathematics, 311-327, Providence, RI: (eds. C. F. Byrnes and C. F. Martin) American Mathematical Society, 1980.
- [59] J. Shamma. *Analysis and Design of Gain Scheduled Control Systems*. PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1988.
- [60] L. M. Silverman. "Inversion of multivariable linear systems," *IEEE Transactions on Automatic Control*, 14(6):270-276 (1969).
- [61] S. N. Singh, M. Steinberg, and R. D. DiGirolamo. "Nonlinear Predictive Control of Feed-back Linearizable Systems and Flight Control System Design," *Journal of Guidance, Control, and Dynamics*, 18(5):1023-1028 (1995).
- [62] J.-J. E. Slotine and W. Li. *Applied Nonlinear Control*. Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [63] S. A. Snell, D. F. Enns, and W. L. Garrard. "Nonlinear Inversion Flight Control for a Supermaneuverable Aircraft," *Journal of Guidance, Control, and Dynamics*, 15(4):976-984 (1992).
- [64] E. D. Sontag. "A Lyapunov-like characterization of asymptotic controllability," *Siam Journal of Control and Optimization*, 13:462-471 (1983).
- [65] E. D. Sontag. "Smooth Stabilization Implies Coprime Factorization," *IEEE Transactions on Automatic Control*, 34:435-443 (1989).
- [66] A. J. van der Schaft. " $L_2$ -gain Analysis of Nonlinear Systems and Nonlinear  $H_\infty$  Control," *IEEE Transactions on Automatic Control*, AC-37:770-784 (1992).
- [67] A. J. van der Schaft. "Nonlinear State Space  $H_\infty$  Control Theory." *Essays on Control: Perspectives in the Theory and its Applications* 153-190, PSCT 14: (eds. H. L. Trentelman and J. C. Willems) Birkhauser, 1993.
- [68] F. Verhulst. *Nonlinear Differential Equations and Dynamical Systems*. Berlin: Springer-Verlag, 1990.
- [69] M. Vidyasagar. *Nonlinear Systems Analysis* (Second Edition). Prentice-Hall, 1993.
- [70] A. Wernli and G. Cook. "Suboptimal Control for the Nonlinear Quadratic Regulator Problem," *Automatica*, 11:75-84 (1975).
- [71] K. Wise, J. Sedwick, and R. Eberhardt. *Nonlinear Control of Missiles*. Final Report, St. Louis, MO: McDonnell Douglas Corporation, Sep 1995.
- [72] K. A. Wise and J. L. Sedwick. "Nonlinear  $H_\infty$  Optimal Control for Agile Missiles," *Journal of Guidance, Control, and Dynamics*, 19(1):157-165 (1996).
- [73] S. Wolfram. *Mathematica, A System for Doing Mathematics by Computer*. Addison-Wesley, 1988.
- [74] F. Wu. *Control of Linear Parameter Varying Systems*. PhD dissertation, University of California, Berkeley, CA, 1995.
- [75] J. Ying-Cun and G. Wei-Bing. "Nonlinear  $H_\infty$ -control and estimation of optimal  $H_\infty$ -gain," *Systems and Control Letters*, 24:321-332 (1995).
- [76] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

## *Vita*

Captain Kelly D. Hammett was born on 20 March 1966, at Wright-Patterson Air Force Base, Ohio. He graduated from Beavercreek High School, Beavercreek, Ohio, in 1984, and attended the University of Oklahoma in Norman on an AFROTC 4-year scholarship. He completed a Bachelor of Science degree in Aerospace Engineering and was commissioned a Second Lieutenant in December, 1988, being named the Outstanding Senior in the College of Engineering and a Distinguished Graduate of AFROTC. Captain Hammett entered active duty in the USAF in January, 1989, attending the Massachusetts Institute of Technology, Cambridge, Massachusetts, as a C.S. Draper Laboratory fellow, specializing in dynamics and control. Earning a Master of Science degree in Aeronautics and Astronautics in June 1991, he was assigned to the Foreign Technology Division (FTD), Wright-Patterson AFB, Ohio, where he worked as a Ballistic Missile Engineer. While at FTD, Captain Hammett performed analysis of metric data on foreign missile tests, enabling USAF weapon system assessments. In August 1993, Captain Hammett attended Squadron Officer School in residence, where he was named a Distinguished Graduate and Outstanding Graduate of his squadron. In Nov 1993, Captain Hammett was assigned to the Flight Dynamics Directorate of Wright Laboratory, also at Wright-Patterson AFB, Ohio. His duties there included research and development on modern aircraft flight control systems, including high angle-of-attack flight enabling technologies. In June of 1994, Captain Hammett entered the Air Force Institute of Technology Graduate School of Engineering, Department of Aeronautics and Astronautics, as a doctoral student, specializing in control and estimation theory. Following graduation in June 1997, Captain Hammett will be assigned to the Phillips Laboratory, Kirtland AFB, in Albuquerque, New Mexico. Captain Hammett is married to Kellie L. (Freeman), formerly of Vian, Oklahoma, and they have a daughter, Taylor, and a son, Chase.

Permanent address: 3400 Billy Lane  
Edmond, OK 73034

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE June 1997		3. REPORT TYPE AND DATES COVERED Doctoral Dissertation
4. TITLE AND SUBTITLE CONTROL OF NONLINEAR SYSTEMS VIA STATE FEEDBACK STATE-DEPENDENT RICCATI EQUATION TECHNIQUES			5. FUNDING NUMBERS	
6. AUTHOR(S) Kelly D. Hammett, Capt, USAF				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology 2750 P Street WPAFB OH 45433-7655			8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/DS/ENY/97-3	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Dr Marc Jacobs AFOSR/NM 110 Duncan Ave, Suite B115 Bolling AFB DC 20332-0001			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved For Public Release; Distribution Unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Nonlinear regulation and nonlinear $H_\infty$ control via state-dependent Riccati equation (SDRE) techniques are considered. Relationships between SDREs and Hamilton-Jacobi/Bellman inequalities/equations are examined, and a necessary condition for existence of solutions involving nonlinear stabilizability is derived. A single additional necessary criterion is given for the SDRE methods to yield the optimal control or guaranteed induced $L_2$ gain properties. Pointwise stabilizability and detectability of factorizations prove necessary and sufficient, respectively, for well-posedness of standard numerical implementations of suboptimal SDRE regulators, but neither proves necessary if analytical solutions are allowed. For scalar analytic systems or those with full rank constant control input matrices, stabilizability and nonsingularity of the state weighting matrix function result in local and global asymptotic stability, respectively, due to equivalence between nonlinear and factored controllability in these cases. A proof of asymptotic stability for sampled data analytic SDRE controllers is also given, but restrictive assumptions make the main utility of these results guidance in choosing appropriate system factorizations. Conditions for exponential stability are also derived. All results are extendable to SDRE nonlinear $H_\infty$ control with additional assumptions. The SDRE theory is illustrated by application to momentum control of a dual-spin satellite and comparison with other current methods.				
14. SUBJECT TERMS Control Theory, Nonlinear Systems, Nonlinear Control State Regulation, Apparent Linearization			15. NUMBER OF PAGES 299	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	